Adaptive, adaptable, and mixed-initiative in interactive systems: An empirical investigation. An empirical investigation to examine the usability issues of using adaptive, adaptable and mixed-iniative approaches in interactive systems.

ADAPTIVE, ADAPTABLE, AND MIXED-INITIATIVE IN INTERACTIVE SYSTEMS: AN EMPIRICAL INVESTIGATION

KHALID AL OMAR

PhD

2009

ADAPTIVE, ADAPTABLE, AND MIXED-INITIATIVE IN INTERACTIVE SYSTEMS: AN EMPIRICAL INVESTIGATION

An empirical investigation to examine the usability issues of using adaptive, adaptable, and mixed-initiative approaches in interactive systems

Khalid Hamad AL OMAR

Submitted for the degree of Doctor of Philosophy in Software Engineering

Supervised by: Prof. Dimitrios I. Rigas

School of Computing, Informatics and Media

University of Bradford

2009

بسم الله الرحمن الرحيم

In the name of God, most compassionate, most merciful

Dedication

This thesis is especially dedicated to the two most precious people in my life, my lovely father and the dearest person on the face of the earth, my soul, my mother. I specially dedicate this work to the person who shares with me the same thoughts and dreams, my soul, my wife. Also, to my lovely sisters and their beautiful children Loloah and Zenah, and to their new lovely babies, Khalid and Hamad.

Khalid Al-Omar

Acknowledgements

I am deeply thankful to Allāh (swt), the most merciful and the most gracious, who makes all things possible, for giving me wisdom, strength and perseverance during this research project and throughout my life.

I would like to give my special heartfelt thanks to my mother, my father, my sisters and my wife, whose unconditional patience, love and support with prayers and advice have enabled me to complete this work.

I would also like to warmly thank my supervisor, Prof. Dimitrios Rigas, for his sincere care, valuable advice and continuous support and encouragement. I have learnt much from him that has richly added to my personality and my academic sense. Thank you.

Many close friends have helped me immensely in different phases through different stages of this research project. I express my gratitude to Mr Abdorhman Abuzaed for his support and help in all stages of this research project. Special thanks go to Mr Emad Alshomrani for his support and help. Also, special thanks go to Nabelah Alsahli for her help with the second experiment. Also, I would like to thank Mr Mehmood Alharbi for his help in the second and third experiments. Thank you all.

Warm thanks go to my friends and colleagues Dr Mohammed Alsuraihi, Dr Mutlaq Alotaibi, Dr Marwan Alseid, Dr Saad Alharbi, Dr Shas Rashid, Dr Fahad Alradadi, Mohamed Sallam, Khalid Ayad, Mansour Al-Shumari, Mansoure Kodari, Yasser Ba Marouf, Turki Alghamdi, Fahad Al-Dosari, Emad Binladn, Naif Alqhtani, Majed Alqhtani, Ali Alshomrani, and Abdullah Alqhtani who supported, encouraged and motivated me during my research work and made this experience pleasant and memorable.

Without these people, as well as others who have gone unnamed, this work would not probably have seen the light of day.

Abstract

This thesis investigates the use of static, adaptive, adaptable and mixed-initiative approaches to the personalisation of content and graphical user interfaces (GUIs). This empirical study consisted of three experimental phases. The first examined the use of static, adaptive, adaptable and mixed-initiative approaches to web content. More specifically, it measured the usability (efficiency, frequency of error occurrence, effectiveness and satisfaction) of an e-commerce website. The experiment was conducted with 60 subjects and was tested empirically by four independent groups (15 subjects each). The second experiment examined the use of adaptive, adaptable and mixed-initiative approaches to GUIs. More specifically, it measured the usability (efficiency, frequency of error occurrence, effectiveness and satisfaction) in GUI control structures (menus). In addition, it investigated empirically the effects of content size on five different personalised menu types. In order to carry out this comparative investigation, two independent experiments were conducted, on small menus (17 items) and large ones (29 items) respectively. The experiment was conducted with 60 subjects and was tested empirically by four independent groups (15 subjects each). The third experiment was conducted with 40 subjects and was tested empirically by four dependent groups (5 subjects each). The aim of the third experiment was to mitigate the drawbacks of the adaptive, adaptable and mixed-initiative approaches, to improve their performance and to increase their usability by using multimodal auditory solutions (speech, earcons and auditory icons). The results indicate that the size of content affects the usability of personalised approaches. In other words, as the size of content increases, so does the need of the adaptive and mixed-initiative approaches, whereas that of the adaptable approach decreases. A set of empirically derived guidelines were also produced to assist designers with the use of adaptive, adaptable and mixed-initiative approaches to web content and GUI control structure.

Tables of Contents

Figures

Tables

Glossary of Abbreviations

ANOVA	Analysis of Variance
HCI	Human Computer Interaction
GUI	Graphical User Interface
E-Business	Electronic Business
E-Commerce	Electronic Commerce
AS	Adaptive split menu
AM	Minimised menu
AH	Adaptive Highlighted menu
MI	Mixed-initiative menu
AD	Adaptable menu

Chapter 1: Introduction to the Thesis

1.1 Introduction

Today, with each new release of software applications there is a plethora of features designed to satisfy every user. As a result, graphical user interfaces have become visually complex and hard to organise and control. This is accompanied by a decrease in the size of the screens of many handheld devices (e.g. mobile phones, PDAs), which further increases the complexity of the interfaces. This complexity has become recognised as a phenomenon which some researchers call creeping featurism [1] and others bloatware [2, 3], it creates conditions where usability problems can arise [3] and where user performance and satisfaction are affected negatively. In response, researchers have sought methods to organise and control such interfaces. McGrenere has suggested multiple interfaces [4] as a solution to software complexity. Others have focused on organising interfaces by using sorting techniques (e.g. alphabetical, numerical, chronological, categorical, or categorical colour-coding) and visualisation techniques (e.g. circular menus [5]) [6]. Both of these approaches are suitable for graphical user interfaces that are easy to organise, while for larger or more complex systems a number of researchers have suggested that it is better to customise the interfaces to the needs of individual users to mitigate their complexity [7], since each user will have different preferences, needs, experience and abilities [6].

Personalisation can be achieved by two contrasting approaches, called adaptable and adaptive, which differ regarding who is responsible for performing the customisation. The adaptive approach dynamically changes the interface layout and content to suit each user's needs, while adaptable interfaces provide customisation techniques which permit users to adjust their layout and content to suit their own needs. Thus, these two approaches differ in their control: adaptive approaches are system controlled, whereas adaptable approaches are user controlled [6].

There has been a debate as to which is the better way to customise interfaces [8], each having its particular advantages and disadvantages, given that by their nature, neither suits the full range of users. For example, adaptable interfaces are user controlled and not all users wish to have full control, for many reasons. For example, they may be busy doing their tasks or simply unable to customise. On the other hand, the main advantage of this approach is that it provides a powerful tool with which users can change and control the system. Conversely, the adaptive approach relies on system control and not all users are willing to relinquish control to the system. The main advantage of this approach is that it does not require much effort from users, while its main disadvantages are lack of control, transparency and predictability. Transparency refers to users being able to understand why changes happen, while predictability means their ability to predict what the system will do. Given these differences, some researchers have suggested a mixed-initiative approach, blending elements of the two approaches to mitigate their disadvantages and increase their advantages [7]. The mixed initiative approach therefore uses both system control and user control at the same time.

There has been some debate in the human-computer interaction (HCI) community as to which of these three approaches is best [8]. One side argues that users should be provided with easily predictable methods to manage their tools, while the other believes that they need the right adaptive algorithm to help them focus on their tasks, rather than on managing their tools [7]. Despite this debate, far too little attention has been paid to comparing the adaptable, adaptive and mixed-initiative approaches.

1.2 The Need for Further Research into Adaptable, Adaptive and Mixed-Initiative Approaches

Recent research on adaptive, adaptable and mixed-initiative interfaces has been motivated by many factors [7], including the following:

- The complexity of software applications is increasing [3, 9].
- There is an increasing problem with information overload when taking the traditional 'one size fits all' approach to software applications, particularly on the World Wide Web [10, 11].
- There is an increasing need to design universal interfaces that match the needs of many users [12, 13].
- The shift from quiet offices to ubiquitous computing increases environmental problems [14-16].
- Screens are becoming smaller [17].

- Despite the continuing debate in the HCI and artificial intelligence research communities as to which of them is best, far too little attention has been paid to comparing the usability of the static, adaptable, adaptive and mixed-initiative approaches.
- Reducing the complexity of visually crowded user interfaces and content is a very important issue and needs further consideration from researchers, since this visual crowdedness increases with each new release of software applications. If the right approach is taken at the right time and with the right user, then the complexity of user interfaces and content will be reduced; if not, the added complexity will annoy users rather than improving functionality.
- The visual channels are becoming very crowded; therefore investigation is required into the use of other senses (e.g. hearing) with the adaptive, adaptable and mixed-initiative approaches to avoid or mitigate their disadvantages and maximise their advantages.
- An examination of the current research on personalisation reveals contradictory findings, including indirect comparisons of static, adaptable, adaptive and mixed-initiative approaches.
- The comprehensive study of usability to examine these approaches has been somewhat neglected in the area of web content and user interface design. Contributors to the research literature tend to focus their studies on examining either performance or effectiveness.

1.3 Aims

The overall aim of this research is to determine which of the adaptable, adaptive and mixed-initiative approaches is best to customise graphical user interfaces and web content. In addition, it aims to understand the factors that make one approach successful at one time and unsuccessful at another. More specifically, it aims to produce a set of empirically derived guidelines for designing more usable personalised interfaces. Furthermore, it aims to identify ways to avoid or mitigate the disadvantages of both adaptable and adaptive interfaces. Satisfying these aims involves answering the following questions:

- Q1: Which of static, adaptive, adaptable and mixed-initiative approaches is more usable in web content and graphical user interfaces?

- Q2: Does the size of content affect the choice of approach to personalisation?

- Q3: How can the disadvantages of both adaptable and adaptive interfaces be mitigated?

1.4 Objectives

In order to fulfil the above aims, we developed two different experimental platforms: an e-commerce website and a set of desktop pull-down menus. The e-commerce platform consisted of four sub-platforms: static, adaptable, adaptive and mixed-initiative. The pull-down menu platform consisted of five different menus of two different sizes: small and large.

1.5 Methodology

The methodologies adopted in this research are:

1. A study of current literature on pure adaptive, adaptable and mixed-initiative applications and on a number of a direct empirical comparisons. The majority of these evaluations compared either (a) adaptive and adaptable, (b) adaptive and static, or (c) static and adaptable interfaces.

2. The data-gathering tools used to assess users' views and behaviour were a post-questionnaire, a questionnaire, pre-interviews, observation and automatic calculation.

3. First Experiment: This was a comparative empirical study that aimed to investigate the usability and controllability of four interactive conditions: static, adaptive, adaptable and mixed-initiative. Each of these was implemented separately as a web-based e-commerce application. The structure of the platform was similar to many web-based e-commerce platforms, except that users could purchase items by clicking or by using one of four types of keyboard: QWERTY, QWERTY with keypad, AZERTY and alphabetical. The differences among the four conditions applied to the

content, position of items on the list and keyboard type. These environments were tested independently by four separate groups of 15 users each.

4. Second Experiment: This investigated empirically the use of five different interactive menu conditions: adaptable, adaptive split, adaptive/adaptable highlighted, adaptive/adaptable minimised and mixed-initiative. The aim was to compare the usability of these five menus, with regard to task accomplishment time and frequency of error-occurrence. The experiment also aimed to examine the effects of different levels of adaptation and adaptability. In addition, it investigated empirically the effects of content size on five different personalised menus. More specifically, it compared the usability of these five types with regard to task accomplishment time and frequency of error-occurrence, effectiveness and satisfaction. In order to carry out this comparative investigation, we conducted two independent experiments, on small menus (17 items) and large ones (29 items) for each of the five conditions. These were then tested dependently using 30 subjects for each size.

5. Third experiment: This investigated empirically the use of improved adaptive, adaptable and mixed-initiative menus. The results of the second experiment had indicated that there were design limitations of these conditions that might affect their usability. Therefore, we utilised the auditory channel for input (speech) and output (earcons). In this experiment users could customise their menus by sound in an incremental manner, following an ‘as you go’ strategy during the tasks. More specifically, the performance of the adaptable menu customised in this way was compared with that of the same menu condition without customisation. This experiment was conducted dependently using 30 subjects.

6. Guidelines: The results obtained from the experimental studies were compared and discussed in order to produce conclusions and empirically derived guidelines for the use of the static, adaptive, adaptable and mixed-initiative approaches.

1.6 Summary of Contribution to Knowledge

To date, far too little attention has been paid to comparing the adaptable, adaptive and mixed-initiative approaches, although there has been some debate in the field of human-computer interaction as to which of these approaches is best [7]. One side argues that users should be able to manage their tools easily, while the other believes that they need the right adaptive algorithms [6]. Despite this debate, research to date has tended to focus on proving one side of the argument, rather than understanding the factors making some of approaches successful in one context and less so in another [8]. The core contributions of this research are summarized below.

1. The primary contribution is to provide an empirical comparison between personalisation approaches in order to assess which condition is best to reduce the complexity of web content and graphical user interfaces. More specifically, it measures the effects on efficiency, effectiveness and satisfaction of static, adaptive, adaptable and mixed-initiative conditions.

2. The second contribution is to provide empirical evidence concerning the effect of menu size on the performance of personalised menus. More specifically, it measures the effect of small vs. large menu size on the performance of adaptive (split, highlighted and minimised), adaptable (adaptable, highlighted and minimised) and mixed-initiative menus.

3. Thirdly, it mitigates the natural limitations of adaptive, adaptable and mixed-initiative approaches by utilising the auditory channel (e.g. speech and earcons) instead of depending on the visual channel only.

1.7 General Thesis Outline

Figure 1 shows the structure of the thesis and how the chapters contribute to the three questions being investigated. Chapter 2 reviews the current literature; Chapter 3 investigates which of static, adaptable, adaptive and mixed-initiative approaches are the best to use in e-commerce; Chapter 4 investigates which of adaptability levels, adaptivity levels and mixed-initiative approaches are the best to use in graphical user interfaces and more specifically in menus; Chapter 5 assesses enhanced adaptable, adaptive and mixed-initiative menus and investigates which of them are the best to use in graphical user interfaces; and Chapter 6 brings the whole work together to provide guidelines to designers and individual users in order to make software more usable. The following paragraphs give an overview of each chapter.

Chapter 1: Introduction to the Thesis

This chapter provides a brief introduction to the research work carried out, the methodology utilised and the contribution of the study to knowledge.

Chapter 2: Literature Review

This chapter reviews all the relevant research in the different areas related to the study aims. In particular, it describes research into personalisation approaches, including a number of direct comparisons of approaches, and multimedia metaphors. The purpose of the review is to help to establish a general understanding and motivation for the study.

Chapter 3: Study One: Comparison of Approaches to Personalisation of Content

This chapter documents the experiment undertaken in order to investigate the usability and controllability of four interactive conditions: static, adaptive, adaptable and mixed-initiative. Each of these was implemented separately as a web-based e-commerce application. The structure of the platform was similar to many such platforms, except that users could purchase items by clicking and/or by using one of four types of keyboard. The differences between the four conditions applied to the content, position of items on the list and keyboard type. These environments were tested independently by four separate groups of 15 users each. Results showed that the mixed-initiative condition was the best in terms of controllability. In addition, surprisingly, subjects who utilised the static condition were found to have a similar level of control to those working under the adaptive condition.

Chapter 4: Study Two: Comparison of Approaches to Personalisation of Graphical User Interfaces

This chapter documents the experiment undertaken in order to investigate empirically the effects of content size on five different personalised menu types: adaptable, adaptive split, adaptive/adaptable highlighted, adaptive/adaptable minimised and mixed-initiative. More specifically, it compared the usability of these five types with regard to task accomplishment time and frequency of error-occurrence. In order to carry out this comparative investigation, we conducted two independent experiments, on small menus (17 items) and large ones (29 items) respectively. These were tested dependently using 30 subjects each. Results show that the adaptable type was surprisingly the most efficient overall of the small menus and the least efficient of the large ones. Conversely, the minimised type was the slowest of the small menus and the fastest of the large ones. Finally, errors were reduced in adaptable and minimised small menus by 50% and 62% respectively, whilst being increased in the large adaptable one.

Chapter 5: Study Three: Comparison of Enhanced Personalisation Approaches

This chapter documents the improvements carried out on menus in order to mitigate the disadvantages of adaptive, adaptable and mixed-initiative conditions. It then reports the testing of the usability of these three menus. Again, two independent experiments were carried out, on small and large menus, this time using 20 subjects each. Results show that the customising menus during the tasks by using speech made task completion faster than not doing so.

Chapter 6: Conclusions and Guidelines

The final chapter provides a concluding discussion of topics including the limitations and contributions of the research. Some areas for further work are also suggested.

In addition, six appendices are provided:

Appendix A: Questionnaire of Experiment One shows the questionnaire used during Experiment One.

Appendix B: Row Data from Experiment One presents the time observations collected during the experiments to measure efficiency.

Appendix C: Results Obtained from T-test for Experiment One provide data analysis techniques have been used to determine statistical significance and to prove the hypotheses of experiment one.

Appendix D: Questionnaire for Experiment Two shows the questionnaire used during Experiment Two.

Appendix E: Row Data from Experiment Two presents the time observations collected during the experiments to measure efficiency.

Appendix F: Results Obtained from T-test for Small and Large Menus of Experiment Two provide data analysis techniques have been used to determine statistical significance and to prove the hypotheses of experiment two.

Appendix G: Questionnaire for Experiment Three shows the questionnaire used during Experiment Three.

Appendix H: Row Data from Experiment Three presents the time observations collected during the experiments to measure efficiency.

Appendix I: Results Obtained by T-test for Small and Large Menus provide data analysis techniques have been used to determine statistical significance and to prove the hypotheses of experiment three.

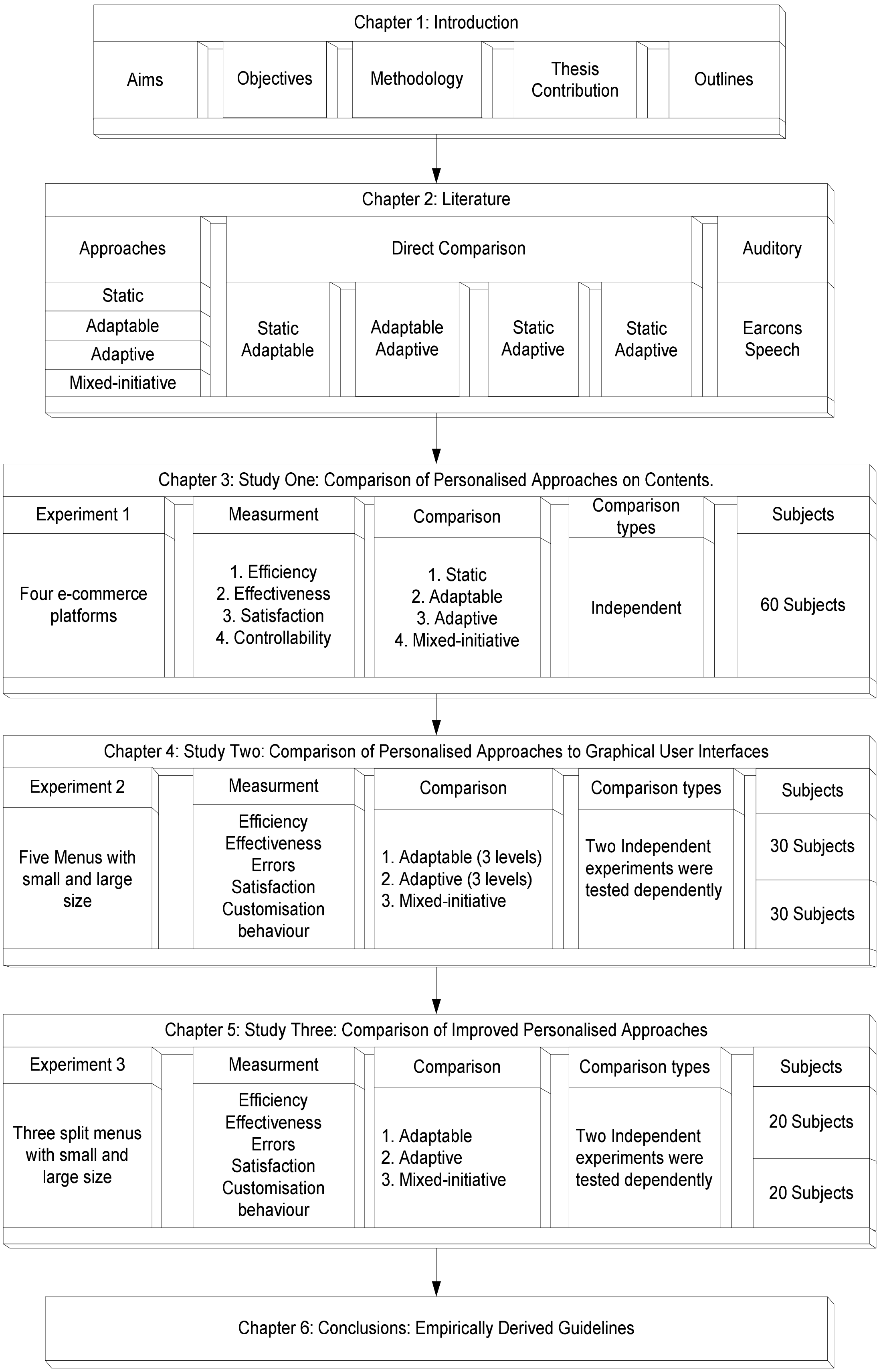


Figure 1: Structure of the thesis and the experimental steps undertaken in the study

Chapter 2: Literature Survey

2.1 Introduction

This chapter describes Human-Computer Interaction (HCI) research related to the personalisation of graphical user interfaces (GUIs) and content. It begins by discussing the problems associated with software applications and their solutions. It then describes the types, levels, and techniques to GUI personalisation. Next, it considers published research related to the three personalisation approaches: adaptive, adaptable, and mixed-initiative. Finally, it makes direct empirical comparisons of the different approaches (see Figure 2).

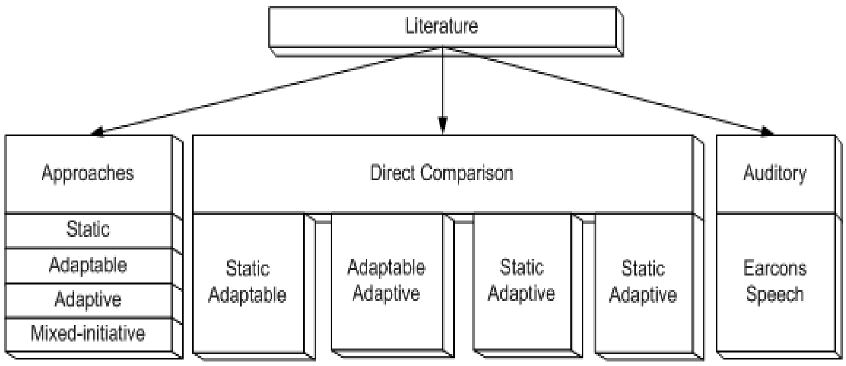


Figure 2: Work related to this study

2.2 Interaction with Graphical User Interfaces and Content

The main goal of HCI research is to improve and enhance the usability of computer systems [18]. The International Standards Organisation (ISO) defines usability as the "the effectiveness, efficiency and satisfaction with which specified users can achieve specified goals in particular environments" [19] (for more details, see [20]). Unfortunately, there are several factors preventing users from achieving their goals by using particular systems.

2.3 Usability Problems

It is becoming increasingly difficult to ignore the growing number of functions in software applications. As a result, the number of menus, icons and toolbars tends to increase, causing interfaces to become visually complex and very hard to organise. There are several reasons for this; for example, users tend to want to own and use the latest technology, and also people like to buy products that have more features [2]. On the other hand, developers produce flexibility by providing multiple functions and styles [21]. In addition, "the designer normally either clutters the screen with many options or builds a structure where users must remember the existence of 'invisible' options and the sequence of actions that lead to them" [22]. Furthermore, the heavy focus on visual interaction and negating the other channels [23]. Using only the visual channel to convey all kinds of information makes complexity worse [24]. This in turn makes users feel nervous, confused [22] and even oppressed [25]. The crowding of interfaces causes users to experience information overload [26, 27] so that they miss important information because "our eyes cannot do everything" [28]. In addition, there are usability problems with visual-only interfaces. These include the complexity of information presented to the user [29], button slipping-off [30, 31], closure [31], interface intrusion into tasks [32] and missing the selection of menu items due to inadequate feedback [33, 34]. For the purpose of this work we shall therefore discuss multimedia and multimodal interfaces.

2.4 Multimedia and Multimodal Interfaces

Today, we are said to be in a multi-channel era, as people can access information through different devices such as telephones, PDAs and televisions [35]. Additionally, information is expanded and brings with it the need to increase the bandwidth of communication between people and machines [36]. Multimodal and multimedia both increase bandwidth [37], but there is an important distinction between them. Users communicate with systems through different channels such as visual and audio, by using different modalities such as visual display, audio, tactile and feedback. Multimodal systems integrate different multimodal prototype; for example, they process two or more combined user input modes [38]. Input modes include speech, pen, touch, manual gestures, gaze, and head and body movements. On the other hand, a multimedia system is one which uses different presentation media (e.g. text, graphics, video and speech) without a commitment to the representation of the information presented [39].

Multimodal choices can differ from user to user and from time to time [35], because users are heterogeneous and their abilities and limitations in this sense are different. Thus, the choice of the interaction mode can be difficult. This suggests a need to personalise interaction modes for each individual user. Meanwhile, multimodal systems have developed rapidly during the last decade [40]. For instance, the MMI2 system [35] is a way of designing computer networks using natural language (English, French or Spanish) through a keyboard, command language, graphics with direct manipulation and mouse gestures. Another system called VoicePaint [41] is a graphics editor application implemented on the Macintosh using Voice Navigator, a word-based speech recogniser board. While drawing a picture with a mouse, the user can talk and ask the system to change the attributes of the graphics context [36]. More recently, a model has been proposed for aircraft cockpit displays via a multimodal interface [42], since the heavy visual workload and physical conditions influence the cognitive processes of an operator in an aircraft. Besides, the command tables in aircraft cockpits impose high loads on human visual senses for displaying flight information such as altitude, vertical speed and airspeed. To date, human-computer interfaces have relied on visual channels to present information to the user [34]. One way of reducing the complexity of a GUI is to reduce the workload on the visual channel. Some researchers have successfully [43] overcome information overloaded by using other sensory modalities such as non-speech sound [44] and haptic technology [43, 45] to provide feedback. This thesis investigates the use of speech as input and output interaction metaphors to enhance usability of personalised approaches. Therefore, for the purpose of the work presented in this thesis, we shall discuss only auditory solutions.

2.4.1 Auditory Solutions (Speech and Non-Speech)

Currently, the use of sound in interfaces has been restricted to providing auditory alerts to users, but many users find that alert sounds (e.g. bleeps) are "distracting and irritating" [46]. There are several other reasons for not using sound, such as noise interference, to which researchers have responded by adapting systems to work within loud environments. Buxton claims that "by effective design, we can reduce the noise component and increase the information-providing potential of sound" [47]. On the other hand, there several reasons for using sound. Gaver argues that "a good first reason to use sound is simply because it's there", he added "sound is more than just an available resource. Sound plays an integral role in our everyday encounters with the world, one which is complementary with vision" [46].

Combining visual and audio channels can significantly improve usability (e.g. [48, 49], [50, 51]). For example, Gaver proposes that "sound should be used in computers as it is in the world, where it conveys information about the nature of sound-producing events". In addition, "using both visual and audio channels would increase the bandwidth of available information" [52]. Rigas suggests that interfaces should use visual metaphors to communicate the information that needs to be conveyed visually and auditory metaphors to communicate other information [23, 29]. Brewster and Crease believe that "sound should become a standard, integrated component in human-computer interfaces" [34]. The results of a study using correlated sound to reduce the time taken to locate visual objects in GUIs showed that time searching for objects visually was significantly reduced when correlated sounds were used [53]. Another study demonstrated that the use of aurally aided search cues significantly enhanced the ability to locate targets and lessened the workload on the visual system in the presence of visual distracters [54].

2.4.1.1 Auditory Icons

Auditory icons are "everyday sounds meant to convey information about events in the computer by analogy with everyday events" [46]. However, they are not without some problems, a major difficulty being to associate environmental sounds with communicated information [55], since the construction of auditory icons assumes natural associations between sounds and events. Some researchers [56] argue that "auditory icons are independent of each other and can only be parameterized with respect to simple object interactions". Nevertheless, auditory icons have been found to be an effective means of communicating information on interfaces. For example, SonicFinder [52] communicates information related to interface icons (bin sound), operations (the copying operation is associated with the sound of pouring liquid) and attributes (object sizes are communicated by the frequency of the sound). Another example is the ARKOLA simulation system [48], which simulates manufacturing plants by conveying sound messages related to ongoing, overlapping events and measures the user's perception of events occurring during the course of production.

2.4.1.2 Earcons

Earcons are structured audio messages [57] which Blattner et al [58] describe as a way of encoding specific information into a sound. There is a growing body of research into the use of non-speech sounds (earcons [58, 59] and auditory icons) to improve the performance and usability of GUIs [27, 54, 60]. Brewster et al conducted detailed investigations of earcons and showed that they are an effective means for communicating information around the interface [61, 62], [63, 64], [56]. Brewster evaluated earcons in order to demonstrate that their use in interfaces could reduce the workload and to provide guidelines [65]. More recently Brewster have derived a guidelines for the presentation of the concurrent earcons to help designer to design interfaces that are more effective at communicating information to users [66]. He suggested that designers should reduce the number of concurrently presented earcons as much as possible. However, they are not without some problems, among the difficulties are that the ad hoc adding of sound by designers who are not sound experts can make the sounds ineffective [67, 68]. In addition, using earcons requires a high level of concentration and the development by users of a perceptual context [69-72], which may lead users to interpret messages incorrectly [73].

Several studies have shown that combinations of speech and auditory icons can significantly improve usability [74]. For example, Rigas et al investigated the use of speech, auditory icons and musical stimuli to convey information to users of a multimedia stock control system [75] and a multimedia email tool [29]. They found that combining these sounds could significantly improve usability [75]. Several other studies have compared different types of auditory representation. For example, one measured the search time and accuracy of menu navigation using four types of auditory representation: speech only, hierarchical earcons, auditory icons and a new type called 'spearcons' [55]. The results show that earcons can provide a much richer sound space than auditory icons.

2.4.2 Speech Recognition

Speech recognition, also known as Automatic Speech Recognition (ASR), is the process of interpreting human speech in a computer [76]. Its main goal is to provide an efficient way of communicating with the computer [76]. The work of Bolt [77] encouraged many others to develop multimodal systems, since using sound can be very beneficial, especially when hands or eyes are equipped [24]. The authors of several studies recommend utilising sound; for instance, Karl & Shneiderman conducted an experiment to assess the utility of using voice commands in parallel with a mouse for word processing applications, against the use of the mouse alone

[78]. They trained 16 users to issue voice commands for 18 tasks such as "page down" and "italic", reporting a 12 to 13% increase in performance. Other examples are given by [79-81]. In addition to personal computer applications, voice commands can be used in environmental control, such as turning lights on and off [76]. Speech recognition is important because it uses a natural human function and because it may be particularly useful for some people such as those who are physically disabled [76]. There are several systems which use sound on their interfaces. For example, InfoSound sound is a composition system which enables users to create and store auditory stimuli and associate them with application events [82]. It has been used to create auditory interfaces for two applications: a telephone network service simulation and a parallel computation simulation. The auditory interfaces helped users to detect multiple event sequences that were difficult to notice visually. Another example is SPHINX, the first accurate, large-vocabulary, continuous, speaker-independent speech recognition system [83]. It was designed to deal with speaker variation, environmental variation and improved speech recognition. Performance improved and the word error rate was reduced significantly by the SPHINX-4 Speech Recognition System [58]. POCKETSPHINX is a more recent version which is a 1000-word vocabulary speech recognition system used for facilitating hand-held devices [84]. However, speech recognition use is not without some difficulties. For instance, speaking while solving a problem is difficult; as Shneiderman points out, "humans can easily speak and walk, but they find it harder to speak and think". In addition, there is a high frequency of errors associated with the limitations of speech-recognition systems. Forsberg [76] enumerates some of the difficulties of ASR, such as noise, ambiguity, speaking style and speed of speech.

2.5 Personalisation

Research suggests that content and interfaces need to provide easy access to the functions that subjects actually use [7]. Furthermore, users tend to use different functions and styles [85] even when performing the same tasks [86]. Thus, "there is no way to organize features in a way that makes essential functionality convenient for everyone" [85]. This suggests the need to personalise GUIs and their content for each individual user [7] and to increase personalisation in many areas of interactive software [87]. Therefore, researchers have sought to improve the usability of interfaces by adapting them to users' needs and by tailoring interfaces and systems to the way people naturally work and live [88]. However, many questions arise in relation to personalisation, such as: Do we need personalisation? Are there any drawbacks to the personalising of systems?

Personalisation [89] refers to the automatic adjustment of information content, structure and presentation to meet the individual needs of users. In other words, personalisation helps to present the right information to the right people [90]. The considerable benefits of personalisation include reduced information overload, while in e-commerce it helps to promote customer loyalty by establishing a one-to-one relationship, increasing both user satisfaction and sales by providing products and services tailored to each individual customer [89, 91]. Information overload on the web often requires filtering mechanisms [92]. Generally, applications that provide a rich functionality usually make systems too complex for some users, while systems that provide a narrower range of functionality are at risk of losing users' motivation. It is often considered that the solution to these problems is personalisation [93]. Furthermore, [8] personalisation helps as user interfaces become more common, many web-users are often less experienced, the number of tasks that users need to complete are continuously increasing, the amount of information available to users is also increasing and users often have limited time. Currently, even search engines may often need to be personalised to allow rapid navigation to very large numbers of products of such variety that it would not be possible to fit all of the material in a printed catalogue [92].

Today, most e-commerce sites welcome personalisation. E-commerce websites are increasingly introducing personalised features in order to build relationships with customers and increase the number of purchases made by each customer [8]. Surprisingly, price is not considered to be one of the top three factors in creating customer loyalty [91], whereas personalisation, is one of the factors that make a customer feel at home. There are many popular personalised websites which help users to manage and personalise their views. The most popular web sites are: Amazon.com, IGoogle, My Yahoo and MyNetscape. For instance, Amazon.com is estimated to use at least 23 different types of personalisation, basing product recommendations on user purchasing history [90] and tracking users' navigation and selection history to produce the "Page that you made". Another example of a commercial site that uses personalisation is Yahoo, which allows users to personalise both the content and presentation of the My Yahoo page [90]. My Yahoo was the first site on the web to use personalisation on a large scale [94]. Manber, Patel and Robison [95] indicate that the Yahoo site applies personalisation in three areas: My Yahoo, Yahoo Companion and Yahoo Search. In My Yahoo, users are able to set up the page to contain information of interest to them and then use this rather than the main entry page to access My Yahoo. The user can choose information of interest from the hundreds of modules that Yahoo provides, such as news, weather and sport. However, such personalisation has drawbacks. According to [95], the discussion of personalisation often raises questions about privacy and security. The threat to the privacy of users is the main drawback when personalisation is applied to software. Many websites collect personal information about users, sometimes without their consent, track their behaviour and build profiles [91].

2.5.1 Categories of Personalisation

There are several ways to provide personalised interactions (also known as customised interactions) to users. Personalisation has focussed largely on information content (such as personalised websites, news delivery and search engine results), and on GUI components (such as buttons, menu items and toolbars). These two categories are affected differently by user expectation and consistency [96]. For example, consistency is not a critical feature of web content, since users expect frequent changes of content, whereas consistency is critical in GUIs because users do not expect changes to occur often. Indeed, consistency is more important than content in the case of GUI components [96]. On the other hand, frequency of usage can influence these two categories, since users will have different frequencies of usage for GUI components and web content. Related work can sometimes be fitted into these two categories, but this is sometimes difficult; hyperlinks, for example, could be considered to belong to both, since they are utilised as control structures in navigation but also have content [96]. A further difficulty arises with customising input devices, for example, as these are not related to either GUIs or content. Despite these difficulties of categorisation, this review will cover the customisation of GUI components and content because there remains an important distinction between these major categories and because we believe that user behaviour will differ from one to the other.

The aspects of content personalisation addressed in this thesis involve tailoring the following:

- Adding, deleting, moving, editing and sorting web content.
- Minimising and maximising the amount of web content.
- Customising lists of displayed items.

On the other hand, GUI personalisation will involve the following:

- Customising GUI component (such as changing menu items positions).
- Minimising and maximising the number of menu items.

2.5.1.1 Personalised Content

There are two processes in personalising content to suit the user: adaptation and presentation [97]. Content adaptation involves deciding what content is most relevant to the user and adapting it before presentation. The strategy for selecting content relies on domain-specific knowledge to different degrees [97]. For example, MASTROcarONTE is an in-car tourist information system providing domain-specific knowledge, since it adapts its recommendations according to user preferences and driving conditions [98]. In contrast, the Generate Evaluative Arguments system [98] bases the generation of arguments on guidelines from argumentation theory; therefore it does not rely on any domain-specific knowledge and can be applied to other domains (see for example [99] and [100]). On the other hand, content presentation involves deciding how to present content to users, such as by the choice of layout, modality and media (see [97] for more details). Content personalisation has been successful in the area of intelligent tutoring systems (such as [101], [102], [103] and [104]), which customise education material to suit users [105], and in the area of recommender systems (such as [106] and [38]), which suggest news stories [107] or products [108] which user might like.

2.5.1.2 Personalised GUIs

The are different ways to personalise a GUI to an individual user [96], for example by adding features such as menu items, toolbar items, buttons or icons to the interface or by using macro mechanisms to produce sequences of functions (e.g. [21, 109], [110]). Another approach is feature management [96], which means customising an existing feature by duplicating its functionality (e.g. [111], [112]), reusing it (e.g. [86]), or managing an existing one (e.g. customising the order [113] or the type and layout [114]). Alternatively, it is possible to customise the appearance of an interface, such as screen colour (e.g. [115], [116]). Feature management and macro mechanisms are both discussed in this thesis, since the former is relevant to GUI complexity, while the latter is more relevant to content complexity.

2.6 Approaches to Personalisation

As noted in Chapter 1, there are three approaches to personalisation: (i) adaptable approaches, providing customisation techniques which permit users to adjust layout and content to suit their needs; (ii) adaptive approaches, where the system dynamically changes the interface layout and content to suit each user's needs on behalf of the user; and (iii) mixed-initiative approaches, which combine these two approaches. We now describe these approaches in some detail.

These techniques all have the same aims, but they are slightly different from each other. Some people find it hard to distinguish between them, especially between the adaptive and adaptable approaches [117], which are not the same, although both seek to customise the content of the system to users' needs and preferences. An adaptable system allows users to change certain system parameters and adapt their behaviour accordingly [88], whereas an adaptive system adapts itself automatically to users based on assumptions about their needs and preferences [88]. According to [118], such a system may adapt the brightness of the user's screen (e.g. according to the stimulus of environmental illumination), but may additionally personalise the colours of the content layout delivered on the screen (e.g. because a specific user suffers from colour blindness).

Thus, these approaches differ in their control of personalisation. Adaptive approaches are system controlled, adaptable approaches are user controlled and mixed-initiative approaches are both system controlled and user controlled at the same time [7]. In addition, there are differences in the techniques they tend to use. For example, adaptive interfaces have tended to use graphical or spatial techniques, or a combination of both, to reduce visual search time [119]. Graphical techniques recognise items and change them graphically, whereas spatial techniques recognise such items and move or copy them for easier access. Adaptive split menus, for example, move the most frequently or recently used items to the top of the menu [120]. An alternative adaptive mechanism which has been introduced recently is that of ephemeral menus, which reduce search time by presenting predicted items immediately, while remaining items gradually fade in [119]. As for adaptable interfaces, these have tended to use coarse-grained or fine-grained methods, or a combination of both, to reduce visual complexity [7]. Coarse-grained techniques include those utilised in the adaptable split menu, allowing subjects to move items to the top or bottom sections of menus, whereas fine-grained ones allow them to move menu items to specific positions [7].

There has been some debate as to which of the above approaches is best [8]. One side argues that subjects should be provided with easily predictable mechanisms to manage their tools, while others are of the opinion that they require the right adaptive algorithm to help them focus on their tasks, rather than on managing their tools [7]. Despite this debate, far too little attention has been paid to conducting a direct empirical comparison of the adaptable, adaptive and mixed-initiative approaches. The following subsections examine in turn each of the three techniques and how it works.

2.6.1 Adaptive Approaches

As early as the 1980s research was directed towards the development of adaptive interaction systems [121]. An adaptive system is one that is capable of monitoring its performance, changing its parameters and improving its performance. Other researchers [122] define it as an interactive system that changes its actions for each user, based on assumptions from information about the user. Overall, most definitions agree that adaptive systems should learn from the actions of each user and then adapt themselves accordingly. During the last decade, researchers have discovered many methods of adaptation and user modelling [88]. First was the pre-web generation of adaptive hypermedia that explored adaptive presentation and adaptive navigation. The second generation was that of the web, which extended the adaptive hypermedia and explored the selection of adaptive content. Finally, the most recent generation is that of the mobile web, which raises new issues in system adaptation, such as location time, platform and bandwidth [88].

2.6.1.1 User Modelling

User modelling involves collecting information about users and distinguishing among them by modelling them. User modelling can be defined as the process of acquiring knowledge about users in order to provide services or information adapted to their specific requirements [123]. The success of systems adaptation is determined by collecting information about individual users [92]. User modelling is an internal representation of user characteristics, such as age, gender, personality, mood and special needs or interests, used by a system as the basis for adaptation.

2.6.1.2 Why User Modelling?

In the absence of user modelling a system will perform in exactly the same way for all users [40]. User modelling is carried out in order to understand users. The problem with this approach is that users all differ in background, knowledge, interests, habits, beliefs and age; and they may have different goals from time to time, or even more than one goal at a time. User modelling can be approached in several ways, such as asking users questions, observing their actions, using stereotypes or a combination of these methods.

2.6.1.3 Collecting Information about Users

There are two ways to collect information about users, Kume [124] suggests: one is to conduct a question-and-answer session requiring the user to state his/her preferences and the other is to monitor the user's dialogue with an application. Kume also indicates that to achieve the ideal configuration via computer-aided adaptation, users should clearly be aware of the user model and must access the information. However, collecting information about individual users presents many challenges, including usability and privacy. Some systems, such as the adaptive tour guide AMPRES [125] and the LifeStyle Finder prototype [126], attempt to collect information about individual users in an indirect way to avoid compromising their privacy [92]. Other systems ask users their opinions about particular topics and use the responses to compare individual users (e.g. Amazon.com [127]). Yet others attempt to record all of the movements and actions that users perform. For instance, researchers have attempted to analyse the navigation actions of users of e-commerce sites to develop a user model, while others have utilised monitoring applications to examine individual interactions with such features as advertisements, promotions, detailed product views, purchases and shopping carts [128]. The main limitation of analysing users' actions in this way is that the data are hard to interpret. Therefore, some systems prefer to ask users direct questions. For instance, during registration, a system may ask users about their interests [92].

2.6.1.4 Levels of Adaptation

There are several dimension of classification of adaptive systems. Two levels of adaptation (technologies) are proposed by [129]: adaptive presentation (at content level) and adaptive navigation (at link level). An alternative classification framework for adaptation is presented by [40], distinguishing layout features such as font sizes, font styles or colours from content. This gives a three-part classification: adaptive presentation, adaptive navigation and adaptive content. The last of these basically means having different content at the level of text, images, videos or animations, from which the system selects the content appropriate to each user depending on his or her profile. Adaptive presentation means changing the layout of the system, such as the interface elements, colours, interface size, font size, font style, image size and number of images, while adaptive navigation means changing the appearance of links, the targets, the number of links on a page or the link type.

2.6.1.5 Adaptation Methods

Adaptation methods are defined [123] as generalisations of existing adaptation techniques, which are defined [40] in turn by a user model representation and an adaptation algorithm. Stephanidis et al [130] document a methodological approach to adaptive systems, pointing out that adaptation is relatively difficult to implement when changes are needed. They focus on the adaptation strategy as a decision-making process which is characterised by the following attributes: what to adapt, when to adapt, why to adapt and how to adapt. There are different methods to achieve adaptation, each of which is based on a clear adaptation idea and can be implemented by using different techniques at a conceptual level.

2.6.1.6 Adaptive Techniques

The adaptation is dynamic, adjusting the interface or the system to support users and help them to cope with system complexity. On the other hand, the personalisation approach relies on each user's direct input [85]. To personalise systems to user needs and preference, techniques are required to hide or show the information, products and services that users may need. There are three popular techniques used to predict users' needs in personalisation: rule-based filtering, collaborative filtering and content-based filtering.

2.6.1.6.1 Rule-based Filtering

A rule-based system allows the administrator to specify rules based on user profiles. These rules then determine the content delivered to the user [90], [91]. One of the more popular techniques in data mining is association rule mining, which looks for items that appear in the data and are associated with other items. The data range from products to a web page that a visitor has accessed. An examples of a rule-based system is [12], [131], [132], [133], [134], [135], [136].

2.6.1.6.2 Collaborative Filtering

Collaborative filtering has been used successfully for recommendation systems (for more details, see [137]). Collaborative filtering systems use information in the form of ratings or preferences and the collaboration of like-minded users to predict as closely as possible certain user preferences [90], [91]. An example of a collaborative filtering system is GroupLens [138], whose clients include Amazon.com [139].

2.6.1.6.3 Content-based Filtering

Content filtering is another technique used for recommendation systems. It works by looking for items that are similar to the items that users liked or purchased in the past. There are several systems available for creating user profiles, included Firefly, whose clients include Yahoo, and Net Perception, whose clients include Amazon.com [140]. Content-based systems rely on content similar to user profiles already obtained [90], [91]. Thus, both collaborative and content filtering rely on user input, allowing them to be combined successfully, for example by Fab [141].

2.6.1.7 Adaptive Content

The most popular and commonly used method for adaptive content is (1) additional explanations, where the goal is to hide irrelevant links and content from the user [123]. This method is used [40] to show additional, prerequisite or comparative explanations. (2) Explanation variants is another method where part of the content page might be hidden or shown depending on user needs and preferences. This method is also known as content variants. The (3) sorting method is used to sort the content of a page. The more relevant the content is to the user, the more it is brought towards the front.

To implement these methods, certain adaptive techniques can be utilised, including conditional text, where all information about a concept is divided into several chunks of text [123], then each chunk of text is associated with a condition. The system will present only the chunk that has a true condition. The stretch-text technique is based on extending the text of the current page, so that the system and the user can extend or collapse the text. This is a very useful technique and can be used to collapse any unwanted or irrelevant content presented to user. In conditional fragments, a user model provides the information that helps the system to determine which chunks of information (fragment variants) should be presented to the user. Page variants is a relatively simple technique which can be applied to the explanation variants method. This technique involves providing two or more alternative presentations of the content of one page. Then, when presenting the page, the system will choose the appropriate content depending on a user stereotype. The most powerful technique of the adaptive content type is the frame-based approach, where information is put into frames, which are associated with rules to hide or show them. For example, Table 1 shows systems that use a combination of methods by applying one or more adaptive content techniques. Many methods and techniques of adaptive presentation and adaptive navigation are listed by [123] and illustrated in Figure 3.

Table 1: Adaptive content: methods, techniques and systems (taken from [123])

Methods	Techniques				
	Conditional text	Stretchtext	Conditional fragments	Page variants	Frame-based approach
Additional explanations	C-Book ITEM/IP Lisp-Critic	MetaDoc KN-AHS PUSH			EPIAIM PUSH
Explanation variants	C-Book		Anatom-Tutor Lisp-Critic WING-MIT	Anatom-Tutor C-Book EPIAIM ORIMUHS SYPROS	Hypadapter
Sorting					EPIAIM Hypadapter

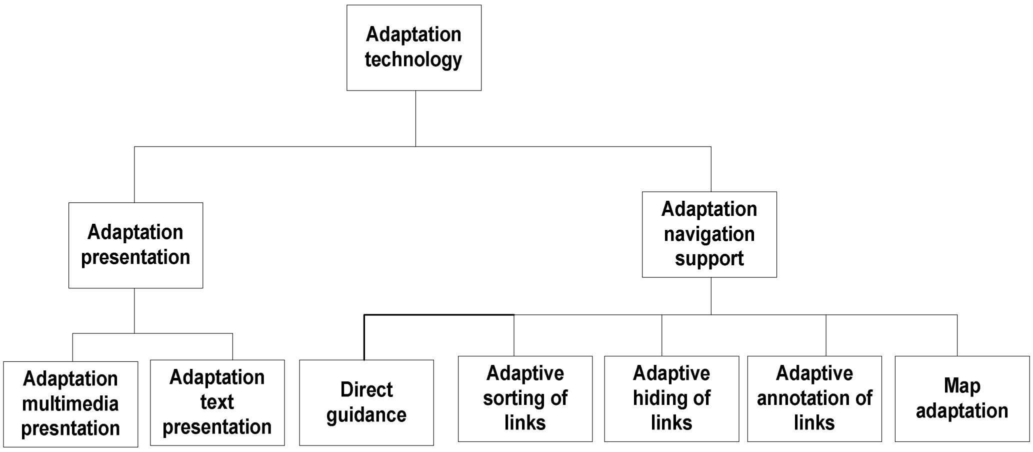


Figure 3: Adaptation technologies in adaptive hypermedia (taken from [123])

2.6.1.8 Adaptive Navigation

The overload of information and pages may increase the need for adaptive navigation in order to help users to find their paths. There are four kinds of link presentation [123]: local non-contextual links, contextual links, links from index and content pages, and links on local or global hyperspace maps. Local non-contextual links are independent from the content of the page, such as a button, a list or a pop-up menu. Contextual links are normally embedded in the content of the page and can be removed. Links from index and content pages are listed as an index and can be consider to comprise a contents page. Finally, links on local maps and on global hyperspace are located in maps which enable users to navigate [123]. Some methods used to support adaptive navigation are [123]:

- Global guidance, whose main objective is to help users to find the shortest navigation path for the required function.

- Local guidance, whose main objective is to help user to find just one navigation step to follow from the current step.

- Global orientation, whose main objective is to improve users' knowledge of the hyperspace in the current position.

- Local orientation, whose main objective is to improve users' knowledge of the different navigation possibilities from the current position and to help them to follow the appropriate links.

Table 2: A classification of adaptation navigation systems according to presentation links and techniques (taken from [123])

Methods	Techniques				
	Direct guidance	Sorting	Hiding	Annotation	Map adaptation
Global guidance	WebWatcher ITEM/IP ISIS-Tutor SHIVA	Adaptive HyperMan CID HYPERFLEX			
Local guidance	Land Use Tutor HyperTutor	Adaptive HyperMan ELM-PE Hypadapter HYPERFLEX	Hypadapter PUSH	ISIS-Tutor ELM-ART	HYPERCASE
Local orientation support (knowledge)		Hypadapter ELM-PE	Clibbon HyperTutor Hypadapter ISIS-Tutor	ELM-ART ISIS-Tutor ITEM/PG Manuel Excel	
Local orientation support (goal)			Hynecosum HyPLAN ISIS-Tutor PUSH SYPROS	ELM-ART ISIS-Tutor	HYPERCASE
Global orientation support			Clibbon Hynecosum HyperTutor ISIS-Tutor SYPROS	ITEM/PG ISIS-Tutor ELM-ART Manuel Excel	HYPERCASE

- Personalised views, which enable users to view and organise hyperspace from a personalised perspective. This requires each user to adjust or adapt the system to his or her personal view.

This technique is subdivided [142] into direct guidance, sorting or adaptive ordering of visible links, adaptive hiding of links, adaptive annotation of links and map adaptation. Table 2 shows adaptive navigation techniques classified into five groups according to the way they are used to adapt the presentation of links [123]. Direct guidance is where the system suggests the next appropriate node or link, depending on the user's goals and needs. This technique can be applied to all kinds of links [123]. The problem is that it provides no support to users who choose not to follow the suggested links. The sorting technique sorts all links on a page according to user need. The more appropriate the link is, the closer it will be to the top. This technique can be applied only to non-contextual links, to indexes and to content pages. The hiding technique is designed to protect users from irrelevant links and content, i.e. those which do not meet the current user's needs and goals. The hiding technique can be utilised with all types of links. Adaptive annotation provides a description of each link in terms of where it will send the user and can be utilised with links of all types.

Finally, map adaptation is a technique applied to a map or a graphical visualisation of the link structure [40].

2.6.1.9 Adaptive Presentation

The goal of adapting the interface is to allow users to use the system more efficiently with minimal error [92]. A well-known example of such adaptation in interfaces is provided by the Smart Menu feature introduced by Microsoft in Windows 2000 [92]. The objective of adaptive presentation is to adapt the layout (font size, colour, language) to the user's needs. Adaptive presentation methods and techniques are usually grouped with those of adaptive content. However, the adaptive presentation is made separately here to give the reader a clear view of it. Consequently, the methods for adaptive presentation are [40]:

- Text presentation such as, whose objective is to adapt the system by providing text in the way preferred by the user.

- Multimedia presentation, whose objective is to adapt all possible presentation and layout features, such as font size, font style and colours, to the user's needs.

2.6.1.10 Empirical Studies of Adaptive Approaches

In contrast to adaptable approaches, where users are responsible for carrying out the customisation, in adaptive approaches, it is the system which customises itself on behalf of the user. According to [143], the first rigorous and successful study of adaptation was reported in 1985, when Greenberg and Witten demonstrated an adaptive interface for a menu-driven application [121], although the long-term effects were not studied. Today, there are several commercial examples of purely adaptive approaches. For instance, the Start menu in Windows XP begins with few shortcuts, then creates additional shortcuts to the most frequently used programs. Another commercial example of an adaptive system is that of Smart Menus, introduced in MSWord 2000. When opened, these first display a reduced version of the full menu which contains the frequently and recently used items, then after a while the Smart Menus will extend and display the full version. When the full version is displayed the position of items is changed, which requires users to scan the menu from top to bottom. Another study showed that more consistency could be attained with a split menu [113], where the menu items are always visible and where the frequently used items are moved to a top section of the menu, separated from other items by a horizontal line. These frequently used items were predetermined at the start of the session and the researchers did not evaluate the top when it changed dynamically according to usage. The experiment, which had 38 participants, showed the benefit of replacing the frequently used menu items at the top of the list. Another example of a split menu can be found in the font selection menus in Microsoft Office, where frequently used fonts are located above a line and all others font below it in alphabetical order. A final commercial example of an adaptive system is Amazon.com, where adaptation is utilised to recommend products to users based on their purchasing history. According to [90], Amazon.com is estimated to use at least 23 different types of personalisation techniques. Amazon.com tracks users' navigation and selection behaviour and uses this to provide a "Page that you made". For registered users, the history will be retained and the system will monitor users' behaviour and clicks.

There have been numerous attempts to evaluate adaptive interface techniques experimentally. For example, a controlled experiment examined two adaptation techniques applied to lists of textual selections [144]. The first was to highlight suggested items by changing the background colour and the second involved shrinking non-suggested items while allowing users to explore these minimised items by means of a virtual fisheye lens. The results showed that accuracy affected the overall user performance and users' ability to locate items that were correctly suggested by the system. Another study examined the effects of predictability and accuracy on the usability of adaptive interface [145]. Their results showed that predictability and accuracy led to improved satisfaction and performance. In a controlled experiment using 26 participants, three adaptive graphical interfaces (split, moving and visual pop-out) were evaluated against a non-adaptive baseline. The authors of this empirical study [40] compared their own results with those of other relevant studies [7, 113, 121] and suggested a number of vital factors that could affect the success of an adaptive interface. These include spatial stability, accuracy and frequency of adaptation, frequency of interaction with the interface and the complexity of the tasks and of the interface itself.

Adaptation has been applied to several application domains, such as intelligent help, information filtering and recommendation, intelligent multimedia systems, education, e-commerce, information retrieval, medicine, tourism, and recently to web services and the mobile web.

2.6.1.10.1 Education Systems

The main benefits of web-based education systems are classroom independence and platform independence [117], while the problem with most such systems is that they consist of a network of static hypertext pages [117], which means that a uniform interaction is provided using the same content to all users despite differences in their knowledge, skills, preferences, and background and ability [142]. Therefore, adaptation becomes important if such systems aim to serve users with diverse needs [117]. In addition, people have different preferences for one or more learning styles. Those who prefer two or more are described as multimodal learners. According to Silverman, there is a theory that “each brain hemisphere is specialized for a different mode of thinking: the left for linguistic, analytical and sequential tasks and the right for artistic, gestalt and creative tasks” [146]. This indicates that there is more than one way of thinking about any particular problem, which means that there are different ways of learning.

Adaptation is intended to overcome one of the barriers to e-learning, which is poor communication between the machine and the learner; however, adaptation techniques differ and it is not always clear which to use to solve a particular problem [147]. The main problems in producing adaptive web-based learning systems [123] are that users have different knowledge and that the knowledge held by any particular user can change quickly. Thus, the same page on a system may be unclear for novice users and boring for advanced learners. Therefore, modern web-based education systems use different types of adaptation techniques [123], including adaptive presentation and adaptive navigation. Adaptive presentation techniques often adapt the content of the page to user knowledge and to the user model [117]. Systems implemented using this technique are [142] the Anatom-Tutor and ITEM/IP. Adaptive navigation support is used to help students to navigate by changing the appearance of visible links depending on the user model [117]. Systems implementing these methods include curriculum sequencing, a technique used to provide the most suitable sequence and to guide the learner through an optimal path.

There are two types of curriculum sequencing [142]: active and passive. Active sequencing involves learning goals and building the best path to achieve user goals, whereas passive sequencing is a reactive technology and does not require any learning goals. This sequence starts when users fail or struggle to solve a problem. Sequencing is divided into high and low levels [142]: the high level sequencing determines the next learning sub-goal and the low level sequencing determines the next learning task. Many current Intelligent Tutoring Systems (ITSs) employ active sequencing [142]. Important components of ITSs are problem-solving support technologies, of which there are three types that can help learners: intelligent analysis of learner's solutions, interactive problem-solving support and example-based problem solving [142]. The intelligent analysis of student solutions deals with the learner's final answer or problem. Interactive problem-solving support helps the learner at each step by providing intelligent help, while example-based problem solving helps the learner by providing examples of similar problems that have been solved earlier. This type was implemented in ELM-ART [117] and ELM-ART II [117, 142].

2.6.1.10.2 Online Information Systems

The goal of online information systems is to provide reference access to information. The problems with such systems [123] include an inability to satisfy users' needs and the fact that users will have different goals in accessing them. They will also often have no time to browse all the information to find the information they are looking for. When a large amount of information is available, users can be divided into those who have a clear goal, so do not need any help with navigation, and novice users who will need such help. Examples of online information systems are Hypadapter [148], HYPERCASE [149], KN-AHS [150], MetaDoc [151], PUSH [152], HYPERFLEX [152], Quality-driven integration of heterogeneous information systems [153] and Adaptive HyperMan [123].

2.6.1.10.3 Information Retrieval and Recommender Systems

Today, Google.com is one of the best known information retrieval systems. Google provides a new service called Personalised Home, which allows users to retrieve personalised information by customising the home page. Users can view the home page they have made via a browser on a PC, on a mobile phone or on a PDA [154].

The main difference between the IR and recommender systems are the criteria utilised in the latter [155]. Among the techniques used to make recommendations and suggestions to users are collaborative [156], content-based [157], demographic [156], utility-based [156] and knowledge-based recommendation [158]. Recommender systems [138, 159, 160] are now utilised in e-commerce sites such as Amazon.com [159].

2.6.1.10.4 Online Help Systems

An online help system is similar to an online information system. The main difference between them is that online help systems are not independent but attached to an application system. Another is that the hyperspace in online help systems is smaller than in online information systems. They share with online help systems the problem of serving different information to different users [123]. Examples of online help systems are EPIAIM [161], HyPLAN [162], Lisp-Critic [163], ORIMUHS [164] and SYPROS [165].

2.6.2 Adaptable Approach

Customisation is often considered to play a key role for organisations that aim to stay ahead of the competition in a global marketplace [166]. Customisation offers users the ability to configure an interface, information or services manually according to their preferences (for example, consider the my.yahoo.com website) [142]. The aims of customisation are to satisfy highly heterogeneous customers' needs at low cost [167], to present customers with very specific products unique to them [168] and to encourage repeat transactions [128]. There are several different ways to achieve this, such as loyalty programmes (special bonuses), the application of proprietary standards for products and services, and by using customisation. Many companies have attempted to engage customers on a more personal level through the processes of mass customisation [167]. However, several research studies have been carried out to determine when and what users can customise. For instance, one study [169] measured the way that different types of customisation affect performance in terms of time taken to complete tasks. Three customisation strategies were considered: 'up front', 'as you go' and no customisation. In the first, users added all interface features before starting the given task, while in the 'as you go' strategy, they customised the relevant interface features of a function at the time that they were required to complete a particular task. In the final case, users did not customise at all [170] but made use of the full interface. The results of this study indicate that the 'up front' strategy was always faster than 'as you go' and customisation was generally worthwhile, particularly for novice users. Thus, the study demonstrated that the most efficient time to customise is before starting a task. Another study compared an adaptable interface to the Smart Menus adaptive interface of MSWord. The results indicate that users usually customised very little, because customisation facilities are often complex and therefore require time, both for learning and for doing the customisation itself [9]. A survey by Fletcher [171] found that 68 percent of web users who personalised an e-commerce site had made a purchase online, compared to 28 percent who had not used personalisation features. However, only 8 percent of the 300 UK websites in Fletcher's survey offered personalisation options to configure the site. Less than one-quarter of the sites in the survey offered registration and only 9 percent made it compulsory. According to [85], most users fail to customise effectively and therefore the challenge is to create improved methods for users to direct interfaces, rearrange functionality and recover from inappropriate adaptations.

2.6.2.1 Adaptation Techniques

Customisation was originally seen mainly as a new paradigm for marketing [172]. There are four approaches as identified by Gilmore and Pine [173] to product customisation: collaborative, adaptive, cosmetic and transparent customisation. The collaborative customisation approach is where companies help users to identify their individual needs and then offer them products which correspond to these needs. The adaptive customisation approach is when the software is designed so that customers can modify it in the absence of any direct interaction with the software company. The collaborative and adaptive customisation approaches thus both involve users in the co-design of the software product. However, cosmetic customisation is mainly concerned with representation and display, rather than the functionality of the software. The transparent customisation approach provides tailored software products without the users being aware that those products have been modified.

Recently, several alternative approaches have been proposed for websites. Among these is the modular product approach [95], which is not very different from other more general product platform approaches for the effective generation of product variants [174]. Other approaches proposed are the generation of recommendations based on user preferences or user similarity compared to other earlier users [93]. Search agents may also help customers to find the products they really need [175], while the creation of several simultaneous product versions designed for different target groups may assist customisation [174]. According to [167], online sites can be customised to users' need and preferences. For instance, each customer can be guided through the purchase process and only the products of interest will be displayed. These sites offer special deals for users who are looking for offers, or provide testimonials to those who want to read what other customers have said about the product.

2.6.2.2 Customisation of Non-software Products

Today, companies offer customisation of products including coffee, cars, shoes, computers and many more. Several leading companies such as Dell, Nike and General Motors now offer customised products. For example, General Motors has developed a concept called AUTOnomy [168] which is enable consumers to customise their cars. Nike offers customised footwear by allowing customers to choose up to 12 different components for each shoe and thousands of colour combinations are possible [176]. Customers can design a personalised name or slogan, and at a small cost even the left and right shoe sizes can differ. All these add to the retail price. Another company that offers customisation is Vans.com, which allows customers to create a personalised version of the company's slip-ons at relatively low cost. The Timberland.com website has a page called "build your own", which allows customers to specify the features of their boots. Polo Ralph Lauren has also introduced the option of customising its products. Thus, many companies offer customisation so that users can contribute to the design of their own products and ensure that no one else will have exactly the same [176].

In addition, companies have begun to customise the entire marketing process, transforming the practice of marketing from seller-centric to buyer-centric [177]. This approach, called customerisation, means customising both the product and the market. Customerisation is more than just mass customisation and is a business strategy to recast a company's marketing and customer interfaces to make them buyer-centric. For example, priceline.com and DealTime.com invite customers to specify the price and a search is then made for companies that are willing to sell at that price [167]. Another company, garden.com, allows customers to design a garden to their own preferences [178], while Dell has established custom websites for their business customers, whose employees can order computer configurations that have been approved by those companies [167].

2.6.2.3 Empirical Studies of Adaptable Approaches

Several studies have been conducted of purely adaptable interaction. For example, Page et al. examined the amount and type of customisation in a study which found that 92% of participants customised their software (WordPerfect 6.0) over a 28-day period [110]. The customisations were minor, such as showing or hiding a ruler bar or creating shortcuts. However, the results showed that customisation had become easy enough to be used often. In particular, it was found that 55% of subjects created shortcuts and 16% created macros. Mackay conducted two other studies (with 51 and 18 participants respectively) to examine the reasons for customisation and to identify the factors that affected it [170]. She found that users might be prevented from customising by factors such as lack of knowledge and being busy, while factors encouraging them to customise included social pressure, software upgrading and external (e.g. job changes) and internal factors (e.g. excess free time). Additionally, she found that 78% of participants had done some sort of customisation. In another experiment, Mackay attempted to distinguish between users who customised and those who did not [179]. There were two groups of users: highly-skilled software engineers and translators, who were less skilled technically. The two groups were located on different research sites, working with two different kinds of customisable software. The first group did not try to determine whether their customisations were useful to other users, while the second group created customisations that were tailored to the needs of others. Another study [109] found that customisation could be affected by the skills of users, who were classified as workers, tinkerers and programmers. They found that workers did not expect to customise, that programmers did expect to customise and that tinkerers lay between the other two groups.

McGrenere et al examined two different strategies in customising their interface: ‘up front’ ’ and ‘as you go’ [169]. They found that 32% of participants customised before the task (up front), while the remaining 68% customised in an incremental manner (as you go). They also found that 63% of participants chose to add all features, while 37% added only the frequently used features. Moreover, no users removed the added features, even if at some point they were no longer needed. This experiment compared two adaptable interfaces to an adaptive one, which we will refer to again in section 2.4.1. Another study presented users with interface façades that allowed quick and simple customisation [180]. In a separate window, users could change widget types and layout, adds features, and then recombines the existing GUI. This study showed that users could interact with multiple interfaces, but it did not show how much they would be willing to use such multiple interfaces.

Other studies have promoted the idea of multilayer interfaces, where users were able to switch between interfaces [94, 181]. This allowed naïve users to start at the first layer and move when needed to a higher one, which they would have to do to enlarge customisation facilities. The design of multiple layers required more careful design, since each layer had to meet users' expectations and needs. Such layered interfaces have been heavily evaluated and used. One example is a study comparing multilayered approaches to a control interface, which found that layered interfaces were better than full interfaces alone in terms of findability, while subjects were more aware of advanced features in the full interface [182]. This study showed that layered interfaces can increase performance as compared to full interfaces, while marked layered ones showed little benefit.

2.6.3 The Mixed-initiative Approach

Fisher [183] discussed in 1993 the need for a system combining the advantages of adaptability and adaptivity. Such systems are known as mixed-initiative systems and were defined in 1999 [184]. To date, far too little attention has been paid to interfaces that combine adaptive and adaptable elements. Today, most adaptive systems are mixed-initiative to some extent [96]. There are some exceptions, including the FlexExcel project [180, 185, 186]; nevertheless, the mixed-initiative approach is common and one use of it is in the dialogue systems, which allow both users and the system to initiate the dialogue. As [96] has pointed out, it is also common in systems that aim to identify the goal of the user in order to assist him in completing his task more efficiently (e.g. [187], [188]). Other areas of use are in recommender systems [38].

There has also been research into such mixed-initiative systems [12]. According to [169], systems should consider users' characteristics and the tasks they need to carry out, prior to the system choosing the right approach for each user. If users are able to customise effectively on their own, then the system will take no action. If not, the system should provide assistance. When the adaptation side applies customisation, it helps to know more about users' preferences. On the other hand, when the customisation system applies adaptation, users are helped to customise more efficiently, such as in [169]. Some systems have adopted the mixed-initiative approach to the content of the system, an example being the Adaptive Education Hypermedia prototype, which produced a system called INSPIRE [189] that was separated into two parts, one being adaptive and the other adaptable. Others have applied the mixed-initiative approach to the interface rather than the content of the system. For example, a system was introduced [185] to provide an environment that adapted the Excel user interface to users and their current tasks. The results suggest that the adaptive component provided potentially beneficial adaptations to users and motivated them to adapt the interface. Another study [169] examined the way that the characteristics of the users' tasks and customisation behaviour affected their performance on those tasks. The results indicated that users may not always be able to customise efficiently and that customisation is beneficial in reducing the time taken by users to complete tasks. In this way, the potential for adaptive support to help users to overcome their difficulties was demonstrated.

2.6.3.1 Categories of Mixed-initiative Approach

According to [96], there are four types of mixed-initiative interaction: conversational interaction, user-controlled adaptation, those where users can override adaptive support and those where users provide relevant feedback. Conversational interaction is common in dialogue systems (e.g. L2Tutor [190] and SMARTedit [191]), where taking the initiative is important to lead the topic. In user-controlled adaptation, users can manipulate the algorithm directly (e.g. Lumiere Project [192]). The user provides relevant feedback and the system is responsible for responding in examples such as

[193], while the user can override the support provided by the system in cases such as [194].

2.6.3.2 Empirical Studies of Mixed Approaches

The FlexExcel system integrates adaptability and adaptivity by providing suggestions to users [180]. The system suggests new menu entries and shortcuts for frequently used items and has a critique feature to provide information to users on customisation. The suggestions and critique feature use a rule-based frequency approach. Once there is a new suggestion, the system notifies the user by sound and with a "tip" icon that blinks three times. FlexExcel was evaluated in an experiment with 13 participants, some of whom were found to have difficulties in initiating the customisation. Another example of a mixed-initiative approach is the Programming by Example system for the HyperCard environment called Eager [195]. This monitors usage and when a pattern is detected a pop-up icon notifies the user that Eager is ready to recommend. When Eager detects a repetitive activity, it highlights menus and objects on the screen. While the user continues to perform the task, Eager anticipates each next action by turning menu items, buttons and text selections to green. Then, once the user is confident that Eager knows how to perform the task correctly, he or she clicks on the Eager icon and the task will be completed automatically. The study found that first-time users were generally able to understand Eager without instruction, but users were uncomfortable with giving up control.

An additional example is a study using adaptive suggestions to reduce the complexity of an adaptive toolbar implemented in MSWord [111]. The system allowed users to add and delete toolbar features. Suggestions were notified by changing the background colour of the toolbar and by using sound. The evaluation of this study is a comparison of adaptable and mixed-initiative approaches; therefore it will be discussed in section 2.8.4. A multi-layer user interface provides users with suggestions on using either a menu-based interface or a command-line interface [196]. The suggestions are based on the number of user errors and users' computer experience. In addition, users are allowed to choose the layer they prefer to use. On the other hand, the system makes some layers visible to users based on their editing history. However, there is no evaluation provided in this study.

A mixed-initiative approach was utilised to support interface customisation implemented to reduce the complexity of graphical user interfaces [105]. The researchers examined two interfaces: a normal MSWord interface and a feature-reduced version of the same interface, containing only features that the user had chosen to add. Their results showed that the mixed-initiative approach was preferred to a purely adaptable approach. In addition, it was found that that the system's suggestions helped to improve task performance. Table 3 compares the adaptive and adaptable approaches.

Table 3: Comparison of Adaptive and Adaptable Approaches

Categories	**Adaptive**	**Adaptable**
Driven by	System [7]	User [7]
Control	System in control [7, 85]	User in control [7, 85]
Adapt based on	System environment; user-provided information and tracking user behaviour [189]	The user's stated Preferences [170]
Changes in content preferences	Change over time based on user browsing patterns, purchases and participation [18]	Do not change unless users update their information [12]
Information about users	Required [87]	Not required [87]
User experience	Not required [91, 95]	Required [95, 112]
Naïve users	Hard to understand [197]	Hard to use [170]
Time of adapting	Not required from users [7]	Requires time, both for learning and for doing the customisation [9]
Users' interests	Users are not familiar to computer-modified and intelligent results; instead they are familiar to static presentation [95]	Users usually customise very little [9] because they do not have time or interest [170]
Complexity	Very complex [9]	Powerful and complex [9]
Users' needs	Confusing [9]	Naïve need help [105]

2.7 Advantages and Disadvantages to Personalisation

The three techniques introduced above (adaptive, adaptable and mixed-initiative), each have many advantages and disadvantages. For instance, adaptive systems have advantages for users such as saving time in locating relevant information, ease of online transaction, and ability to connect quickly with the right resources and people [38]. In addition, adaptation provides some unique benefits to an organisation such as real-time data about member preferences, ability to classify member data to identify interests in various categories, ability to deliver targeted content and recommendations to individuals and organisations to help to plan programmes, products and services that meet the needs of members. On the other hand, there are some disadvantages to users and organisations [38]. Adaptive approaches provides less control to users, as they might miss content, and there might also be privacy concerns [38]. The disadvantages to organisations are that expectations may be raised about adaptive condition to solving more basic web usability problems, that personalisation and customisation costs will vary and that they will require resources. Adaptive approach may also limit the opportunities to cross-sell [38].

Manber, Patel and Robison [95] believe that for high-skill personalisation, usability is the most difficult technical issue. They give the issue of predictability as an example to demonstrate the weakness of personalisation. Also, they believe that scalability must be built into any web personalisation, because people expect the software to interact with users quickly. However, they point out that My Yahoo users usually do not customise [169, 170], but accept the facilities offered on the default personal page (in fact, this increases the need for adaptation). They suggest that users' failure to customise is because the default page is good, because the customisation tools are difficult to use and because most people do not need complex personalisation. In addition, they indicate that users are not accustomed to unexpected, surprising and intelligent results, but are used to static presentations. In general, they argue, users do not understand customisation, therefore it is very important to present customisation in an intuitive way, not as something surprising to users.

Table 4: Advantages and disadvantages of adaptive and adaptable approaches

	Adaptive	Adaptable
advantages	▪ Reduced complexity of software, control, and help to avoid overloading or "underloading" the user [94]. ▪ Time-savings in locating relevant information [38]. ▪ Increased speed and accuracy [40, 94]. ▪ Enhanced user learning process, decreased search and navigation time, and reduced risk of being lost in hyperspace [40]. ▪ Ability to connect with the right resources quickly, provision of products & services that meet needs [38].	▪ Time-savings in locating relevant information [38]. ▪ Ease of doing business online, and ability to connect with the right resources and people quickly [38]. ▪ Customisation facilities are often powerful [9],
Disadvantages	▪ Lack of control over the process, lack of transparency, and lack of predictability [169]. ▪ User registration required, Profile required to be built, collecting personal data, and complexity of implementation [91]. ▪ Data and privacy protection problems [21]. ▪ User is observed by the system, and user can be distracted from the task [21]. ▪ Complexity of implementation, and difficulties in evaluating adaptive systems [94].	▪ Users usually customise very little [9]. ▪ Customisation complex, and require time both for learning and for doing the customisation [9]. ▪ Users lack interest and time, and difficulty in modifying settings [170].

Mackey [170] studied the customisation activity of 51 staff members using a UNIX software environment and found that when users faced a problem they could either change their behaviour or customise the software. From the users' perspective the second choice took more time but created fewer risks. The study also showed that there were many triggers of and barriers to customisation, categorised as external events, social pressures, software changes and internal factors. Examples of factors triggering people to customise were job changes, office moves, going on trips, breakdowns and upgrading, while the most common barriers to customising were lack of time and interest, and difficulty in modifying settings. Mackey found a solution to shortage of time by allowing users to share their customisations.

Adaptation benefits human-computer interaction by increasing speed and accuracy, enhancing the user learning process, reducing search and navigation time, limiting the risk of being lost in hyperspace, which often causes systems to lose users, and helping to improve user satisfaction [1], [40]. The effects of different kinds of adaptive systems [89] could differ according to the area of application and the perspective [123]. However, there is one drawback to applying adaptation to any system, which is that the interface is less stable for the user. This results in confusion, especially for naïve users. From the implementation point of view, adaptation is difficult to achieve because adaptive systems are complex and expensive [40]. The main advantages and disadvantages of both adaptive and adaptable systems are listed in Table 4.

2.8 Direct Empirical Comparisons of Personalised Approaches

Many researchers have sought to reduce selection time by making recently and frequently selected items easier to choose. An examination of the current research on personalisation reveals contradictory findings. For example, in a controlled experiment, 26 subjects were asked to search for names in a telephone directory accessible through a hierarchy of menus and this was tested against a static system. Subjects performed faster with the adaptive system, which 69% of them preferred. In addition, results showed that the adaptive system reduced the search paths for repeated names, reduced time per selection by 35% and reduced errors per menu by 40%. Trevellyan and Browne [147] replicated this experiment with a larger number of trials because they believed that subjects would eventually become familiar with the static menu and memorise the required sequence of key-presses. They found that the adaptive system was effective and that after using it for a long period of time users did begin to perform better with the static interface. Another study compared an adaptive menu with a static one. In a controlled experiment, sixty-three subjects were requested randomly to complete 24 tasks using both menus. The results showed that the static menu was faster than the adaptive menu on the first group of tasks, while there was no difference in the second group of tasks between the static and dynamic menus, because subjects in both groups were able to increase their performance significantly. Eighty-one percent of the subjects preferred the static to the adaptive menu [166]. Another example, is a static interface was compared to three adaptive alternatives as follows: (1) split interface, where important functions were copied into an extra toolbar; (2) moving interface, where important functions were moved into a toolbar and (3) visual popout interface, where important functions were moved and made visually prominent. Two experiments were conducted. The first had 26 participants and investigated the impact of the different interfaces under two adaptive algorithms (frequency vs. recency based). The results showed little difference between the interfaces for the cognitively more complex task, while on the less complex one, the split and moving adaptive interfaces were faster than the static interface. Furthermore, in terms of satisfaction, perceived benefit and perceived cost, the split and moving adaptive interfaces were found most beneficial and least costly, and they were preferred in the more complex task. The visual popout interfaces were found distracting. In the less complex task, there was less support for the adaptive interfaces. The second experiment was conducted with 8 participants and compared adaptation accuracy (70% vs. 30%). The results showed that user performance worsened as the adaptive algorithm's accuracy decreased. Another between-subjects study with 40 participants examined an adaptive approach to command line usage [142]. It compared (1) a command-line interface, (2) a menu-based interface, (3) a hybrid interface, where participants had access to both the menus and the command line, and (4) an adaptive interface, where the system moved users from the menus to the command line. It was found that the adaptive interface was significantly faster than the non-adaptive, hybrid approach. Another study compared the performance of adaptive and static menus [123]. More recently, a study examined a new adaptive technique called ephemeral adaptation. Ephemeral menus present predicted items immediately, while remaining items gradually fade in [122]. These new techniques were examined with static and highlighted adaptive menus. The results showed that ephemeral menus were faster and preferred over the static control condition when adaptive accuracy was high, and no slower when adaptive accuracy was low. In addition, ephemeral menus were faster than highlighted adaptive menus, while both were preferred to static menus.

2.8.1 Static vs. Dynamic

There are several experiments that have taken place to compare static and dynamic applications and interfaces. For example, an experiment [7] was carried out with 27 subjects to compare the performance (speed and error rate) of static, adaptive and adaptable menus. Each menu was implemented separately. It was found that the static menus were faster than the adaptive ones, but not faster than the adaptable menus unless the subjects used the adaptable menu first. The adaptable menus were faster than the adaptive ones, except when subjects used adaptable menus first. It was found that 55% preferred the adaptable menus, 30% the adaptive menus and only 15% the static ones. This indicates that a strong majority of users wanted a personalised interface. It was also found that the adaptive menus were slower than the adaptable ones except when subjects used the adaptable menu first [7]. This study demonstrated some other important issues. First, ease of use is not sufficient in customisation, because some users did not recognise the value of customisation; therefore the authors suggest that users should be guided by being provided with examples. They also suggest that providing users with control will lead to a better perceived performance and higher overall satisfaction. This study demonstrates that users can customise effectively. Other work by McGrenere showed that users are willing to customise their menus [105].

2.8.2 Adaptive vs. Adaptable

In addition to the comparison between static and dynamic techniques, research studies have compared the adaptable and adaptive techniques. Direct comparisons of adaptive and adaptable approaches have also had conflicting results. For example, a six-week field study with 20 participants evaluated two interfaces combined with adaptive menus in the commercial word processor MSWord 2000. These were a personalised interface containing desired features only and a default interface with all the features. During the first four weeks of the study participants used the adaptable interface, then used the adaptive interface for the remaining time. It was found that 65% of them preferred the adaptable interface, 15% favoured the adaptive interface and the remaining 20% chose the MSWord 2000 interface. However, according to [85], there were two potentially confusing variables. First, MSWord 2000 and the proposed interfaces had very different designs, which may have differed in their usability. Second, all participants completed the adaptive condition after the adaptable condition. In another study, McGrenere et al. [7]carried out a controlled laboratory experiment with 27 participants to compare the efficiency of three of the Sears and Schneiderman [113] split menus. The first of these was a static split menu, the second an adaptable split menu where the top half was adaptable by the user and the third an adaptive split menu, where the system would dynamically assign the top half based on frequency and recency of selection. The experiments found no interactive effect between order and menu. On the other hand, the comparison was complicated, according to [85], because performance depended on menu order and subjects were exposed to the three conditions, although when they were not presented with the adaptable interface they were significantly faster with the adaptive or static ones. The findings were that split static menus were significantly faster than adaptive menus. The adaptable menu was faster than the adaptive menu when participants were guided by example, because they were able to understand the value of customisation. In addition, results showed that in these circumstances there was no significant difference between the adaptable and static menus. Nevertheless, 55% of subjects preferred the adaptable menu, 30% the adaptive and 15% the static. In another laboratory experiment with 18 participants, Jameson and Schwarzkopf [198] directly compared automatic recommendations, controlled updating of suggestions and a condition where no recommendations were available. The comparison was concerned with content rather than the graphical user interface. In the automatic recommendation (i.e. adaptive) system, the updating was performed automatically by the system, while in the (adaptable) system using controlled updating of recommendations, it was done by users and in the third (static) system, no recommendations were provided to users and the system did not change during usage. Jameson and Schwarzkopf found no differences in performance score among the three conditions.

2.8.4 Mixed-initiative vs. Adaptable

Most studies in the field of personalisation have been limited to the differences and similarities among the static, adaptive and adaptable approaches. Consequently, there has been a small amount of research into mixed-initiative interfaces. Very few references were found in the literature to direct comparisons of a mixed-initiative system with either an adaptive or an adaptable alternative. One of these rare studies compared a mixed-initiative toolbar with adaptable one. Specifically, it compared an adaptive bar (mixed-initiative system) with the built-in toolbar present in MSWord (adaptable system) [111]. It found that the mixed-initiative system significantly improved performance in one of two experimental tasks. In another study, Burnt et al. [96] designed and implemented the Mixed-Initiative Customisation Assistance (MICA) system, which provided subjects with the ability to customise their interfaces according to their needs, while also providing them with system-controlled adaptive support. They found that users preferred mixed-initiative support and that the MICA system's recommendations improved time on tasks and decreased customisation time.

2.9 Critical Assessment

Many researchers have attempted to reduce the complexity of GUIs and content by using adaptive or adaptable approaches, each of which has its unique challenges. For example, lack of control, predictability, transparency, privacy and trust are the main issues in adaptive interfaces, whereas in the adaptable approach to customisation, time, difficulty and lack of interest are the main difficulties [170]. Recently, several researchers have attempted to evaluate the effects of these drawbacks, for example, by evaluating personalisation approaches using different levels of controllability [9, 105, 111]. In addition, several studies have attempted to overcome the limitations of these approach. For example, some researchers have suggested multiple interfaces as a solution [4], while others have attempted to use both approaches in such a way that one could be used to support the other [105, 169, 199]. Some researchers have suggested allowing users to overrule any adaptation actions, while others propose recommending adaptations to users and leaving them to make the final decision to accept or reject these suggestions [186]. However, little research has been done on interfaces and content that combine both adaptive and adaptable approaches (i.e. mixed-initiative). Furthermore, an examination of the current research into adaptive, adaptable and mixed-initiative approaches indicates that researchers seem to have neglected other channels, such as sound and haptics. An equally small amount of research has examined the factors making one personalisation approach successful one time and unsuccessful another time [143]. This includes studies examining the effect of screen size in adaptive and adaptable approaches [17, 200]. Finally, very little work has been done to evaluate directly and empirically the adaptive, adaptable and mixed-initiative approaches to graphical user interfaces or to content. This research attempts to fill these gaps in the literature. Consequently, the first experiment, reported in Chapter 3, examines the application of these approaches to content, while Chapter 4 reports an experiment to examine the use of personalisation techniques in graphical user interfaces.

2.10 Summary

This chapter has presented a review of the literature related to the personalisation of content and interfaces. First, most of the approaches that exist today are either purely adaptive or adaptable. Second, there is no work discussing in detail the balance that users need in mixed-initiative systems between adaptive and adaptable elements.

However, different studies have been reviewed in order to show how effective these approaches are in solving the problem of bloatware. For example, direct comparisons have been considered between adaptive and adaptable approaches, static and adaptive ones, and mixed-initiative and adaptable ones. In addition, the use of sound (auditory icons and earcons) and speech have been reviewed in order to show how effective these approaches are in solving the problem of creeping featurism. The results of previous studies of the adaptive and adaptable approaches have shown a conflict as to which approach is most able to reduce the complexity of software applications in user interface and web content. While there are many studies which show the strengths and weaknesses of these two approaches, there has been very little investigation on the use of speech recognition to reduce the visual information overload or solve the usability problems of GUIs and content. In addition, little is known on the impact on performance and user satisfaction of customising by using speech recognition; nor has personalising GUIs and content by different multimodal approaches been evaluated. There is a need to investigate thoroughly the use of sound as input and output in personalisation, since there is evidence that combining different sense (such as visual and auditory) could reduce significantly the complexity of both GUIs and content. Therefore, there is a strong need to investigate the potential of mitigating the drawbacks of adaptive and adaptable approaches by utilising other communication channels such as speech in order to enhance usability and provide a new set of empirically derived guidelines.

Far too little attention has been paid to dealing with the complexity of graphical user interfaces and content through mixed-initiative approaches. There has been no work focusing on evaluating in depth the mixed-initiative approach to GUI and content. The exception to this is one direct comparison of a mixed-initiative approach to GUI customisation with either the adaptive or adaptable alternatives [96]. However, there is a suggestion made by researchers that mixed-initiative techniques can improve performance. For instance, [190, 191, 201]. The following chapter attempts to answer the first part of question one presented in relation to the main aim of this thesis (see Section 1.3). Therefore, it sets out the hypotheses used to conduct the first empirical investigation in this work in order to investigate which of the personalisation approaches (adaptive, adaptable and mixed-initiative) to web content (such as e-commerce) users prefer and why.

Chapter 3: Study One: Comparison of Personalised Approaches to Content.

3.1 Introduction

This chapter documents the initial experiment carried out to investigate which of the approaches to the personalisation of menus (adaptive, adaptable or mixed-initiative) is the most usable in terms of efficiency, effectiveness and satisfaction. For purposes of comparison, the traditional (static) approach was also assessed. The systems were built as web-based e-commerce applications. The main reasons for choosing this domain are the growing debate between adaptability and adaptivity in web-based e-commerce applications. Furthermore, the interfaces in this environment are by nature visually complex and can have unbounded catalogues. Consequently, using a standard search and browsing facility becomes difficult, increasing the need to adopt one of the alternative approaches. This is particularly important when subjects are obliged to interact with the systems provided. The empirical evaluation of the initial experiment is expected to provide a preliminary understanding of how much controllability subjects would prefer (see Figure 4).

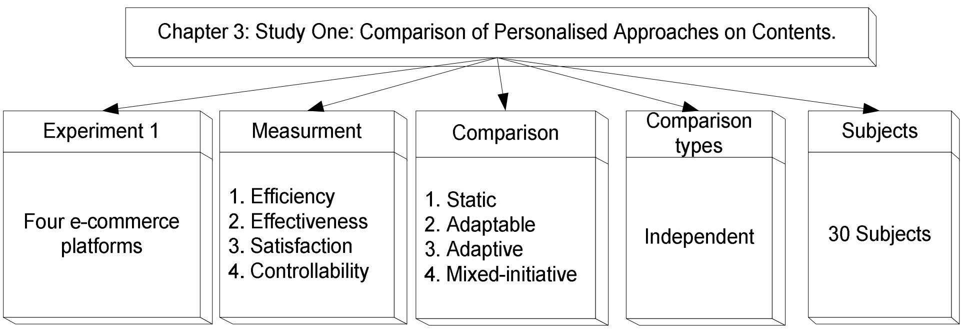


Figure 4: Summary of chapter 3

3.2 Problem and Solutions

It is becoming increasingly difficult to ignore the growing number of functions in software applications. As a result, the number of menus, icons and toolbars tends to increase, causing interfaces to become visually complex and very hard to organise. The visual complexity of interfaces has become recognised as a phenomenon which some researchers call creeping featurism [202] and others bloatware [203]. To overcome this problem and reduce their visual complexity, interfaces need to provide easy access to the functions that subjects actually use. Some studies of interfaces have focused on organising menus by using sorting techniques such as alphabetical order and categorical colour-coding [5]. Other studies have focused on visualisation. For example, circular menus have been developed so that all menu items are equally distant [5]. However, these are temporary solutions where interfaces become visually more complex. More than ever before, subjects need to simplify them to suit their individual requirements. This indicates a need to personalise interfaces in some way [7].

There are two main approaches to meeting this need: in the adaptive approach, interfaces and their contents are dynamically modified to match each user's needs, while adaptable interfaces provide customisation techniques which permit subjects to adjust layout and contents themselves. A third approach to personalisation is to combine the adaptive and adaptable methods in mixed-initiative interfaces [120]. These approaches differ in their control of personalisation. Adaptive approaches are system controlled, adaptable approaches are user controlled and mixed initiative approaches are both system and user controlled at the same time [7]. In addition, there are differences in the techniques they tend to use. For example, adaptive interfaces have tended to use graphical or spatial techniques, or a combination of both, to reduce visual search time [119]. Graphical techniques recognise items and change them graphically, whereas spatial techniques recognise such items and move or copy them for easier access. Adaptive split menus, for example, move the most frequently/or recently used items to the top of the menu [120]. Moreover, ephemeral menus have recently been introduced to reduce search time by presenting predicted items immediately, while remaining items gradually fade in [119]. On the other hand, adaptable interfaces have tended to use coarse-grained or fine-grained components, or a combination of both, to reduce visual complexity [7]. Coarse-grained menus allow subjects to move items to the top or bottom, whereas fine-grained ones allow them to move items to specific positions in the list. For example, these techniques are utilised in the adaptable split menu to allow subjects to move items to the top or bottom part [7].

There has been some debate as to which of these approaches is best [8]. One side puts forward the view that subjects should be provided with easily predictable mechanisms to manage their tools, while others are of the opinion that they require the right adaptive algorithm to help them focus on their tasks, rather than on managing their tools. Despite this debate, far too little attention has been paid to conducting a direct empirical comparison of the static, adaptable, adaptive and mixed-initiative approaches.

3.3 Aims and Objectives

The aim of the initial experiment was to shine new light on this debate through an examination of the usability of static, adaptive, adaptable and mixed-initiative menus. More specifically, it aimed to investigate the significant differences among these approaches in terms of efficiency, effectiveness, learnability and satisfaction. In addition, the initial experiment aimed to elicit subjects' views of the level of control provided in each condition.

In order to fulfil these aims, three objectives had to be attained. The first was to measure precisely the efficiency of each condition by timing the completion of tasks and counting the number of clicks, pages visited and errors in each condition. The second objective was to measure the effectiveness of each condition by calculating the percentage of tasks completed successfully by all subjects and the number of subjects who successfully completed all tasks. The third objective was to obtain the subjects' views on ease of use, ease of purchasing, ease of navigation, ease of shopping and overall satisfaction. It was also part of this objective to obtain subjects' assessment of the level of control provided by each condition and the level of control required by each subject.

3.4 Experimental Platform

The experimental platform was a typical web-based e-commerce application. For example, subjects had to register first to log in, then they could purchase items and view their basket before proceeding to payment. More specifically, each platform consisted of a different type of page such as registration, login, view basket and assist. Each platform also contained a menu and keyboards. It was decided to implement a typical web-based e-commerce application to examine how subjects would interact with such a system and to explore how interaction metaphors affected the search time and effort. The experimental platform utilised four types of interaction conditions: static, adaptable, adaptive and mixed-initiative approaches (see Figure 5 and 6). Each of the four conditions was implemented separately and all applied principally to the contents (list of items), keyboards and layout.

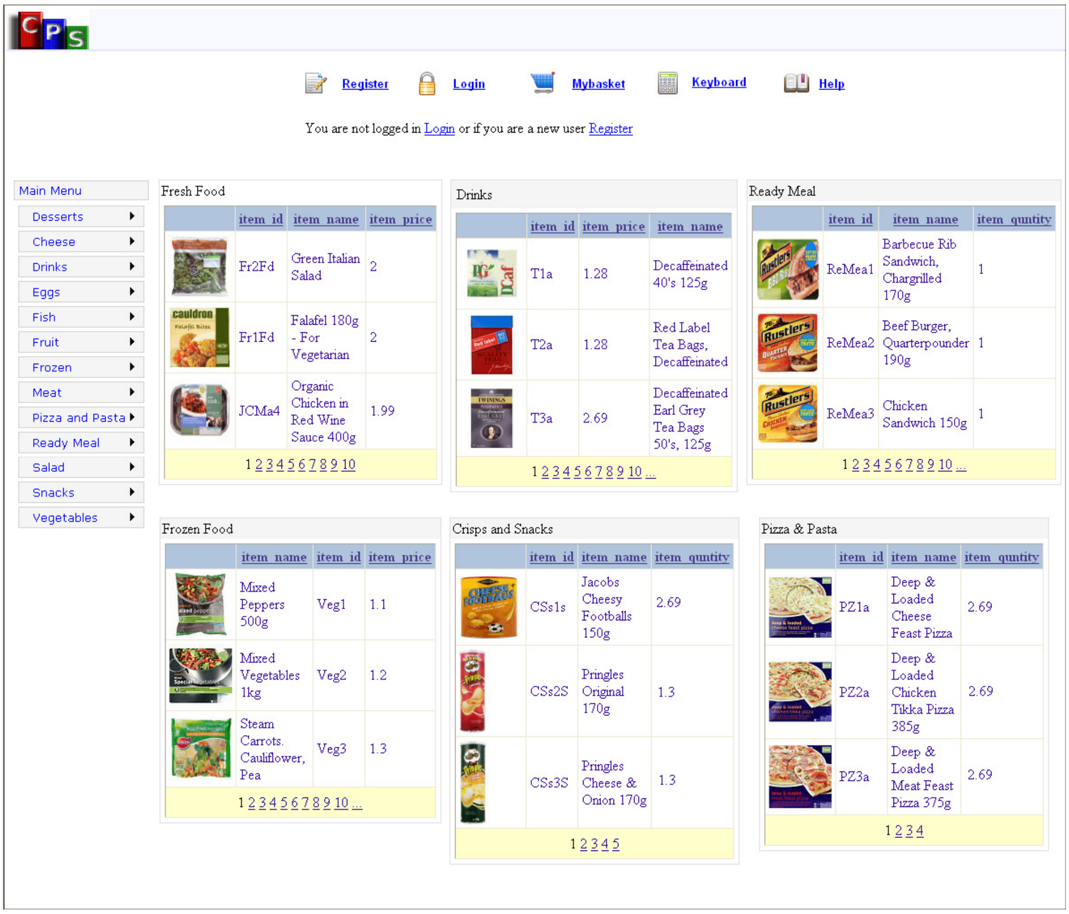
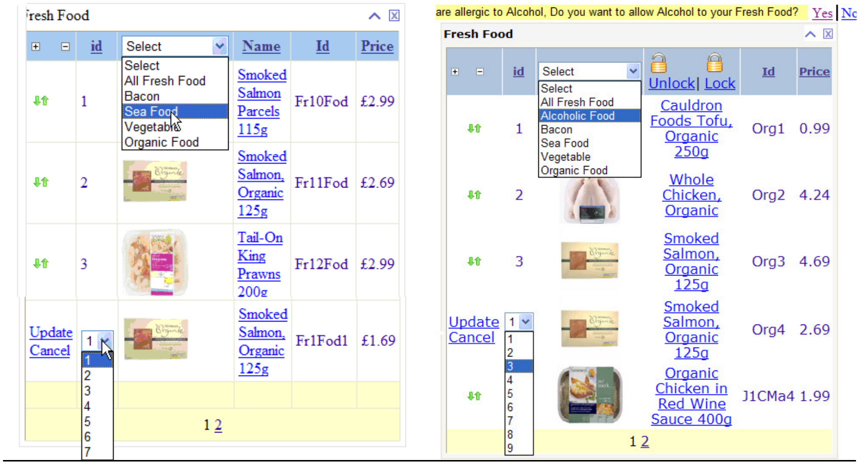


Figure 5: Main User Interface

(a) Static content (b) Adaptive content



(c) Adaptable content (d) Mixed-initiative content

Figure 6: Types of items list

3.4.1 Contents (item list)

Items on the main page were grouped in six categories, each consisting of 10 to 50 items. The same amount of information was displayed for each category. More specifically, this information comprised the name, ID, picture and price of each item in the category. The default number of items displayed at the beginning of the experiment was four per group. The other items were hidden and subjects had to search for the required item within each category. Groups in the static condition did not change during use by subjects, whereas in the adaptive condition, after each selection the selected item would move to the top of the list, then the system would count how many times each item had been used, accept the first item and update the list. On the other hand, in the adaptable and mixed-initiative conditions, subjects were allowed to add a new list to the main page and delete an existing list.

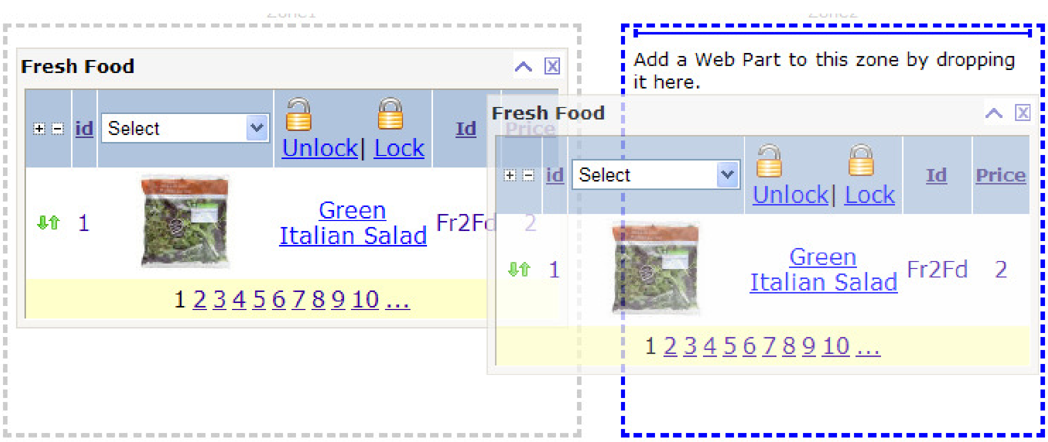
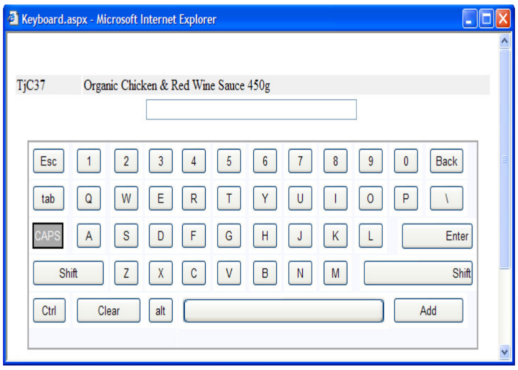


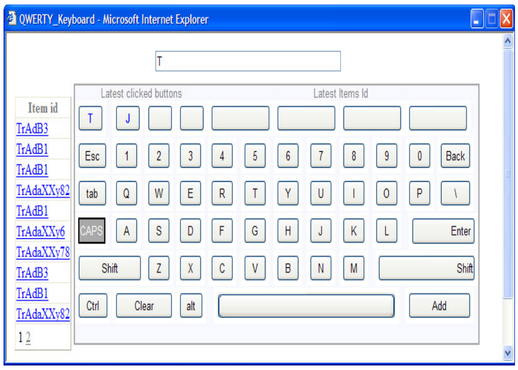
Figure 7: Changing list position

They were also able to change list positions by dragging and dropping lists from one zone to another (See Figure 7), to move items to a specific location on the list (up or down) and to customise the number of items displayed (from 1 to 10) in each category. In the mixed-initiative approach, subjects could additionally lock a list to prevent items from moving up or down, or unlock one which had been locked. Finally, if subjects attempted to add non-personalised items, the list would warn them by displaying a 'confirm' message.

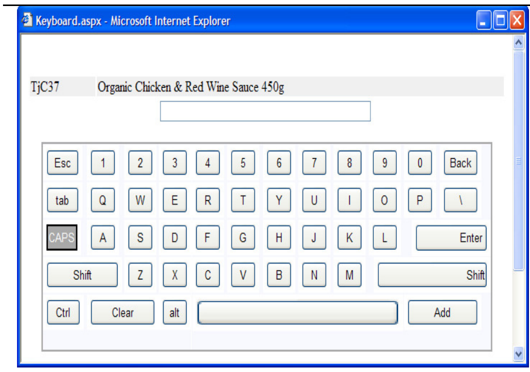
3.4.2 Personal Keyboard

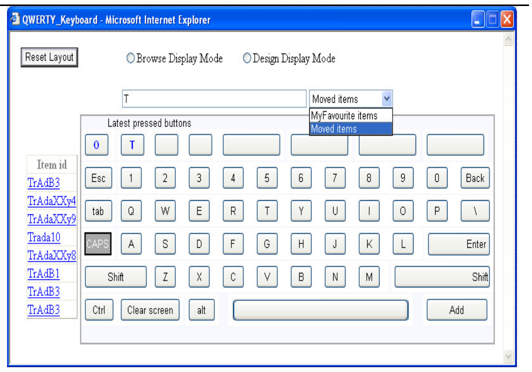
The structure of the platform was similar to many web-based e-commerce applications, except that subjects could purchase items by using either a mouse or a keyboard (see Figure 8). The whole platform, including the keyboard, was personalised to the user. In other words, there was a static keyboard in the static environment, an adaptive keyboard in the adaptive environment, an adaptable keyboard in the adaptable environment and a mixed-initiative keyboard in the mixed-initiative environment. In order to personalise the different approaches to the keyboard, an automatic text completion feature was utilised. This displayed alphabetically 10 items at a time to help users to select the desired item. However, automatic text completion was different for each keyboard, since the environments were different. For example, in the static condition, all items were auto-completed, since no changes occurred in item lists, while in the adaptive condition, only those items that had been customised by the system would be auto-completed. In the adaptable condition, only those items that had been customised by the user would be auto-completed, whereas in the mixed-initiative condition, both personalised and customised items would be auto-completed. On the other hand, all keyboards captured the last four buttons pressed and item IDs entered, displaying these at the top of the keyboard, except that in the adaptable condition, it was the user's responsibility to display themTherefore, an automatic text completion In order to purchase by keyboard, subjects had to enter (by mouse clicks) the item ID. Four types of keyboard were developed: QWERTY, QWERTY with keypad, AZERTY and alphabetical. Each condition integrated different keyboard schemes. In the static condition, the QWERTY keyboard was provided as the most familiar type. In the





(a) Static keyboard (b) Adaptive keyboard





(c) Adaptable keyboard (d) Mixed-initiative keyboard

Figure 8: Four types of keyboard

adaptive condition, subjects could choose only one of the four types of keyboard before starting the experiment. In the adaptable and mixed-initiative conditions, the four types of keyboard were provided together and subjects could switch from one to another at any time. However, in the mixed-initiative condition, the QWERTY keyboard was suggested to subjects as the default type.

3.4.3 Static platform

For the static platform, the contents, layout and keyboard did not change during the course of use. The goal was to design the ideal platform to do the required tasks as efficiently as possible. In order to do this, the content was used according to predetermined tasks and placed on the main page. The QWERTY keyboard was chosen as being the standard keyboard that most subjects were familiar with. Thus, the content and the keyboard were considered ideal for carrying out the tasks (See Figure 6 (a)).

3.4.4 Adaptive platform

In the case of the adaptive platform, the layout, contents and keyboard changed automatically during use. The goal was to design a predictable system, personalised as much as possible. Therefore, subjects were asked before using the interface to indicate which type of keyboard they preferred and to choose some new contents based on our scenarios. For example, if a subject selected an organic food as a preferred item, then all organic foods in the list were placed at the top. However, when the participant started, four items were displayed as a default in each web part on the home page. The order of items in the list was then changed according to the subject's selections by means of two algorithms, taking account of frequently and recently used items. These were adopted as being the two most popular algorithms, used by Microsoft and suggested by the literature (Findlater and McGrenere, 2004 [7]). Thus, after each selection the software counted how many times each item had been used and updated the list (See Figure 6 (b)). The adaptive keyboard provided an automatic text completion function. This would auto-complete only items that matched user preferences and those which had been purchased before (See Figure 8).

3.4.5 Adaptable platform

In the adaptable platform, the layout, contents and keyboard were changed by subjects before and during use. The goal was to make the customisation process as easy as possible. Therefore, we provided two levels of customisation for subjects to modify the lists of items: coarse-grained and fine-grained [7] techniques were available for users to move items to a specific location (See Figure 9). The main page provided two choices for the user: either a vacant page where the user could decide freely which content to add, or a suggested full set of contents. This was done because some of the early studies suggested the need to examine full-featured versus reduced interfaces. Then, when the participant started, four items were displayed as a default in each web part of the home page. Subjects were able to customise the display with as many items as they liked (with a minimum of one item). They could also sort the web contents by item name, ID and price, and search in different subcategories. Based on the scenario, on the main home page the system displayed all items. In addition, the system allowed subjects to add new content to the home page and move items inside the list. Thus, changing the contents of the home page was entirely left to each user's responsibility.

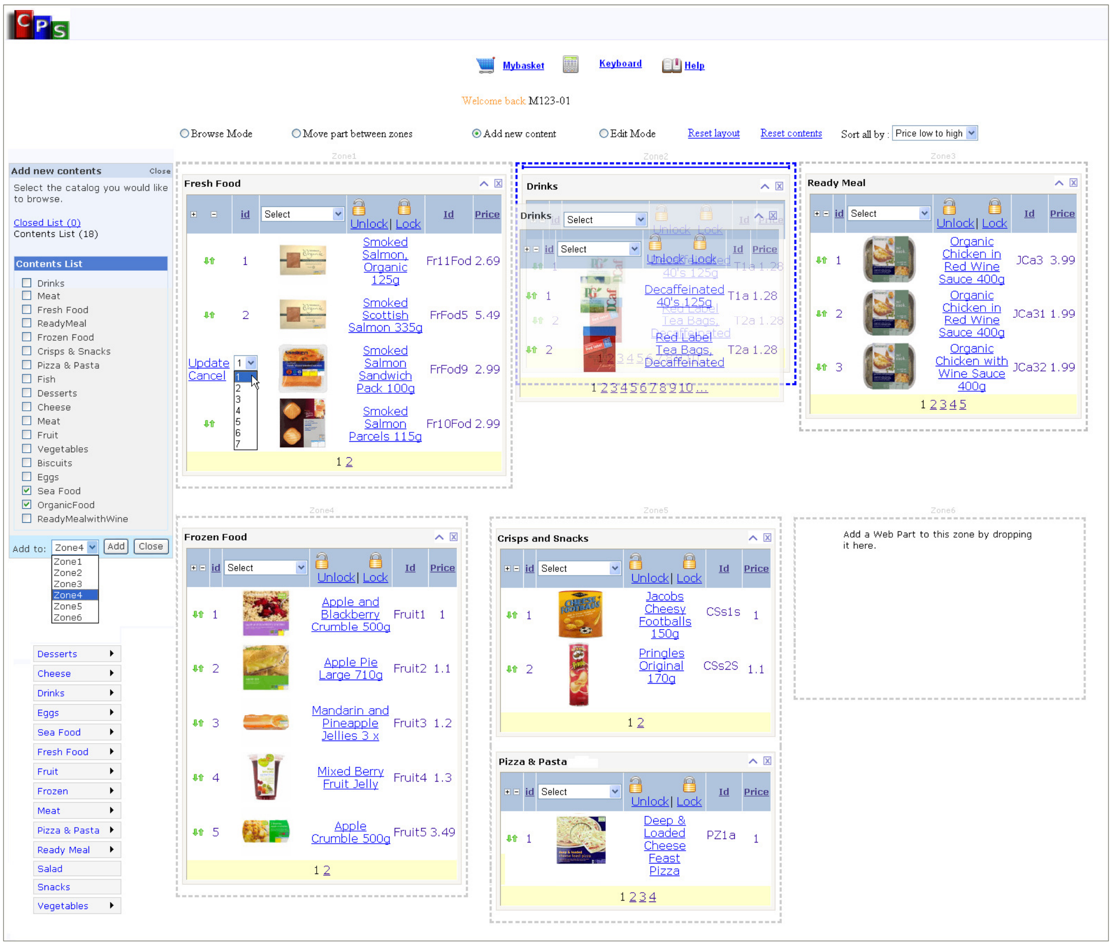


Figure 9: Adaptable Interface

3.4.6 Mixed-initiative Platform

In the mixed-initiative condition, control was shared and the goal was to ensure that it was shared as fairly as possible. The mixed-initiative algorithm was dynamically determined, based on the most frequently and recently used items. However, to allow subjects to take control, a new function was implemented to lock and unlock item movement (Figure 6 (d)). Items were moved to the top of the list when clicked three times, even if the list was locked. Initially, when the website was loaded, the default content of the home page was personalised. Thereafter, the user was responsible for organising and locking the lists. Keyboard auto-completion worked with all personalised items as well as those which had been customised, which assisted subjects with both types of item. Based on the scenario, the system adapted the content of the main home page by not displaying items from the Fruit, Eggs or Alcohol lists. In addition, the system allowed subjects to customise the home page by adding or deleting contents and moving items within lists. In this condition, the system did not display organic items at the top of the list and users were responsible for customising the lists. In the mixed-initiative condition, users had to choose which of the four keyboards would be the default before starting the experiment. In other words, all four keyboards would be available for users to switch from one to another, but they were required to select their preferred keyboard as default.

3.5 Experimental Hypotheses

The aim of Study One was to measure the usability (efficiency, effectiveness and satisfaction), and controllability of static, adaptive, adaptable and mixed-initiative menus in e-commerce. Based on the literature review of related work, the following hypotheses were made, to be tested by the study.

H1: The adaptive, adaptable and mixed-initiative approaches will be more efficient than the static approach in terms of task accomplishment time, frequency of clicks and pages visited, frequency of error-occurrence, number of tasks completed successfully, and preferences.

H2: The adaptive approach will be more efficient than the adaptable approach in terms of task accomplishment time, frequency of clicks and pages visited, frequency of error-occurrence, number of tasks completed successfully, and preferences.

H3: The mixed-initiative approach will be more efficient than both the adaptive and adaptable approaches in terms of task accomplishment time, frequency of clicks and pages visited, frequency of error-occurrence, number of tasks completed successfully, and preferences.

H4: The adaptable and mixed-initiative approaches will be more adequate than both the static and adaptive approaches in terms of controllability.

3.6 Experimental Methods

3.6.1 Subjects

The sixty subjects from the general population (forty-four males and sixteen females) who completed Study One were divided into four independent groups of fifteen each for the empirical work. Subjects in Study One were divided into four independent groups of fifteen each for the empirical work, since the experiment had four independent conditions. Therefore, participants were randomly assigned to one group each, in order to mitigate the learning effect that might otherwise occur. Therefore, we decided to have 15 subjects for each condition because we felt that in an initial comparison, this number would provide us with vital indications of the benefits and drawbacks of each approach, at the same time as keeping the experiment under control. Subjects in all groups were asked to accomplish the same group of tasks (three easy, three moderately difficult and three difficult), as well as one learnable task before starting each level of tasks. Each user attended a five minute training session about the environment before doing the requested tasks. All subjects were between the ages of 18 and 40; 70% of them were postgraduate students. Most used the internet for 10 hours or more each week. A large majority (85%) stated that they did not customise new software unless they had to, while the remaining 15% stated that they did so. A third of subjects (19 subjects) had never used any customisable web pages, while 57% (34 subjects) had done so once and just 11% (seven subjects) used these every time they went online.

3.6.2 Experimental design

The hypotheses listed above were tested empirically using a between-subjects design (i.e. each subject participated in only one condition) (see Figure 10). This design was considered ideal for Study One because each condition was designed to last approximately two hours, so there would have been a significant learning effect if a within-subject design had been used. Each subject was assigned randomly to one of the four groups and so to a condition and set of tasks.

3.6.3 Tasks

All subjects were asked to accomplish the same group of tasks and one learnable task before starting each group. The training tasks were provided to assist subjects in learning how to perform the main tasks. Subjects were informed that they were training tasks. The main tasks were designed at three levels of complexity: easy, medium and difficult. In order to avoid the impact of the learning effect, the order of complexity was varied between subjects. To ensure a variety of complexity, a design guideline was followed. More specifically, the number of available items, position in the list, number of requirements and guidance were considered when designing the tasks (Table 5). In the easy tasks, subject searched within a list comprising a

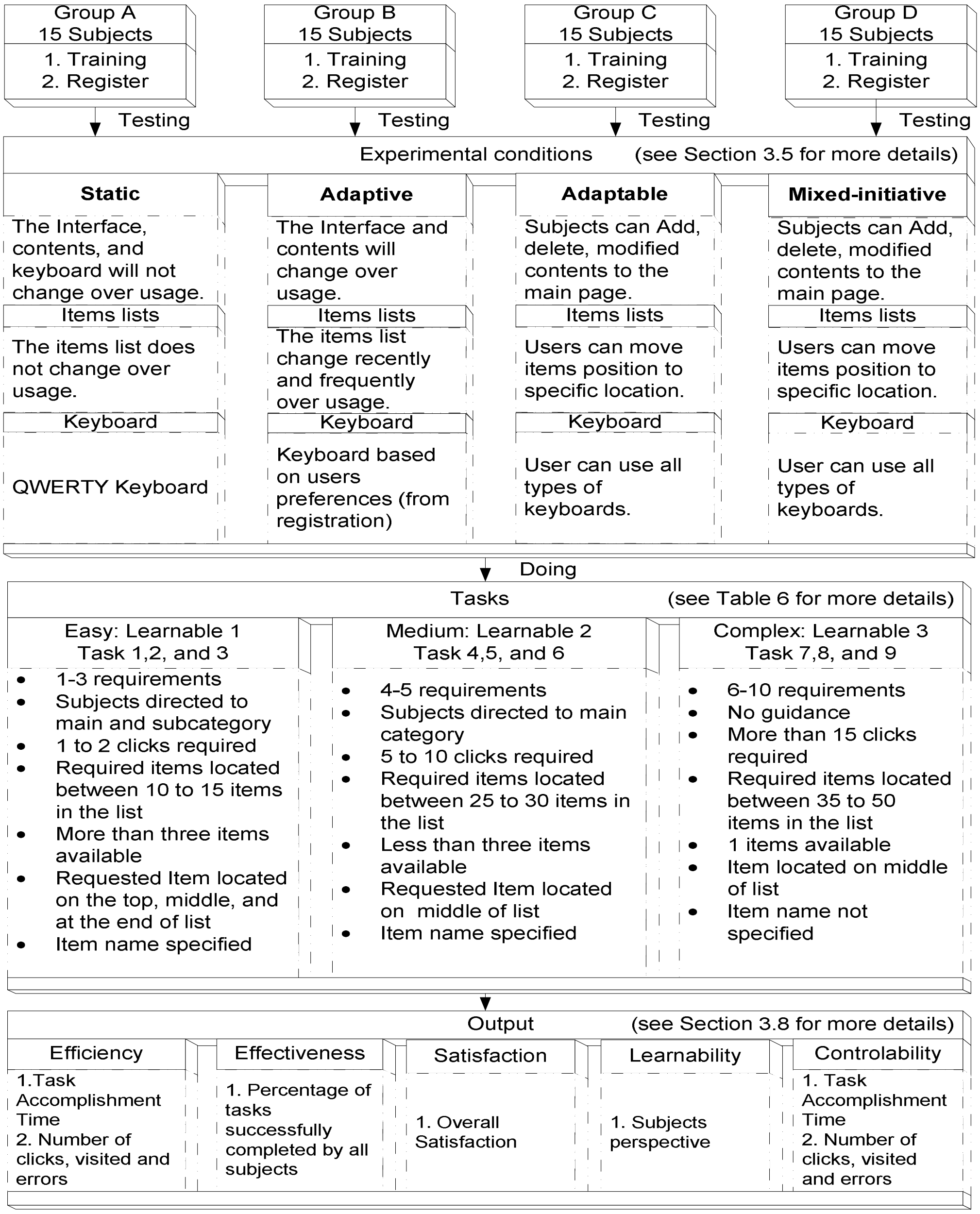


Figure 10: Plan of Study One

maximum of 20 items, where the required item was placed at the top, middle and at the end of the list. There were fewer than four requirements and subjects were guided by the provision of the name of the list and the subcategory. In the case of medium tasks, the number of items on the list was increased to 30 and availability was reduced to two items. The required item was placed in the middle of the list, there were between four and six requirements and subjects were guided to the list but not to the subcategory. Finally, for the difficult tasks there was only one item available within a list of more than 40 items. Items were again positioned in the middle of the list, there were more than seven requirements and no guidance was provided.

Table 5: Task Design

Category	Easy Tasks	Medium Tasks	Complex Tasks
Number of requirements	1-3	4-5	6-10
Guidance Type	Subjects directed to main and subcategory	Subjects directed to main category	No guidance
Number of clicks required	1 to 2	5 to 10	More than 15
Number of pages visited	None required	Maximum of 2	More than 5
Number of items in the list	10 to 15	25 to 30	35 to 50
Number of items available	More than three	Fewer than three	Only one
Item name specified	Yes	Yes	No
Item position in list	Top, middle & end	Middle	Middle

3.6.4 Independent and Dependent Variables

Independent variables are those which were controlled during the experiment to ensure its consistency. They were:

1. Tasks: All subjects had exactly the same number of tasks (3 learnable and 12 main) at the same levels (easy, medium and complex). This was ensured by following criteria developed to ensure the consistency of tasks (Table 5).

2. Interactive metaphors: All subjects within the same group were required to assess the same set of interaction metaphors.

3. Learning effect: to ensure that the learning effect was controlled in Study One, the subjects were assigned randomly to one environment and to different groups of tasks.

4. Task criterion time: Each task had a criterion time within which it had to be completed and would be regarded as unsuccessfully completed if not completed within this time.

5. Amount of training: Training was recorded to ensure that all subjects within the same group had the same amount of training. In addition, we explained the same

amount of information to all subjects in all groups. For example, all groups were shown how to purchase by using the keyboards. However, some groups had extra training time and information, depending on the environment they were testing. For example, the group testing the adaptable approach had extra information about how they could customise the interface.

6. Assist: To ensure that all subjects had the same amount of assistance when they needed it, an 'Assist' section was provided in the main interface of each environment. Subjects were also allowed to ask any questions before and after each individual task.

The dependent variables, which were those which we measured test the hypotheses, were grouped according to their matrices.

Efficiency

1. Task accomplishment time: The time taken to learn how to do a particular task plus the time taken to complete it.
2. Number of clicks and pages visited: The number of clicks required and pages visited in carrying out each task.
3. Number of errors: The number of errors made during each task. Errors were counted during performance of the main tasks but not during learning tasks.

Effectiveness

1. Percentage of tasks successfully completed by all subjects: The number of tasks correctly completed within their criterion times as a percentage of total tasks performed.
2. Number of subjects who successfully completed all tasks: The number of subjects who correctly completed all tasks within their criterion times.

Satisfaction

1. Overall satisfaction: Subjects' satisfaction was measured for each interaction metaphor utilised during the experiments.

Controllability

1. Subjects' opinions were obtained about the level of control each condition provided and the level of control they needed.

Customisation behaviour

1. Customisation time: The amount of time spent on customisation.
2. Added and deleted items: The total number of items added to and deleted from the interface.
3. Locked and unlocked items: The total number of locked and unlocked lists.

3.6.5 Procedure

The set of tasks was designed to fit into a forty-five-minute session. The experimental procedure was as follows. (1) Before the experiment a questionnaire was used to obtain information on subjects' demographic factors and on their computer and customisation experience. (2) Subjects were given a 5-minute tutorial on using the system and to explain the benefit of the approach used. (3) Before each group of tasks, a scenario was provided, along with a practical learnable task, to allow subjects to familiarise themselves with the approach. Subject were told to ask questions if they needed to, regarding the environment that they were evaluating or the experimental procedure. (4) At the end of each session subjects were asked to give ratings for the environment tested. The performance of each user was observed, recorded and noted in an evaluation form. (5) After each group of tasks, subjects were allowed a short break before completing a questionnaire giving their views about the tasks and the approach. For the adaptable approach, subjects were encouraged to customise but informed that they had the right not to do so. They were invited to customise before starting the experiment and at any time they felt the need. In addition, instructions for customisation were given and assistance provided to subjects when needed. For the adaptive approach, subjects were asked to register with the system before starting the experiment. Instructions for registration were given and assistance provided when subjects needed it. Finally, for the mixed-initiative approach, subjects were asked to register with the system first and then to customise it after reading the experimental scenario.

3.6.6 Training

Each subject attended a five-minute recorded training session about their environment before attempting the tasks. Further explanation was also provided when needed.

3.6.7 Data Collection

Quantitative and qualitative data was collected by recording the experiments and from questionnaires, interviews, observations and written notes. Experiments were recorded to provide detailed data on subjects' performance and to document any errors that occurred. It also allowed precise calculations of task completion times, frequency of clicks and pages visited. Subjects were not told that the experiments were being recorded, to ensure that they would perform the tasks without any distraction. The questionnaires and interviews provided qualitative data from subjects' perspectives on matters such as their satisfaction. Recording observations and taking notes during the experiments helped to provided full understanding of each condition and to collect the required data. All of these measures are described below, grouped according to category.

3.6.8 Measurement

In order to fulfil the aim of the study, three objectives had to be attained. The first was to measure precisely the efficiency of each condition by timing the completion of tasks and counting the number of clicks, pages visited and errors in each condition. The second objective was to measure the effectiveness of each condition by calculating the percentage of tasks completed successfully by all subjects and the number of subjects who successfully completed all tasks. The third objective was to obtain the subjects' assessments of ease of use, of purchasing, of navigation, of shopping and of their own overall satisfaction, as well as their views concerning the level of control provided by each condition and the level of control needed.

Efficiency can be measured in terms of effort required to accomplish a goal or task [204, 205]. Here, it was measured by the time taken to complete tasks, the number of mouse clicks and pages visited, and the number of errors made in doing each task. Effectiveness can be measured in terms of whether goals are met or tasks completed successfully [204, 205]. Here, it was calculated as the percentage of subjects who completed (learning and main) tasks and as the percentage of tasks completed by all subjects. To compare effectiveness across the four conditions a critical time for task completion was derived for each level of difficulty (easy, medium and complex). Thus, a task would be regarded as successfully completed if subjects finished it within the critical time. Satisfaction and controllability are usually considered subjectively and were measured qualitatively in this experiment by attitude rating scales asking subjects to rate their satisfaction with and control over each interaction condition [204, 205]. The metrics and dependent variables are set out in Table 6. One-way ANOVA was utilised to examine the differences among all groups in order to identify any differences in efficiency, effectiveness and user satisfaction. Testing for significant differences between means was important for this research, as its purpose was to discover which condition was most usable.

Table 6: Metrics and dependent variables

Metrics	Dependent variable
Efficiency	1.Time taken to complete the tasks 2.Number of mouse clicks and visited pages 3.Number of errors
Effectiveness	1.Percentage of Tasks successfully completed 2.Number of Subjects who successfully completed all Tasks
Satisfaction	1. Overall Satisfaction
Learnability	1.Time taken to complete the training tasks
Controllability	1. Subjects Perspective

3.7 Results

This section considers the experimental results obtained from both qualitative and quantitative measures, of self-reported and observed data. In addition, interviews were conducted with subjects when needed. It was noticeable that subjects who participated in the evaluation of the adaptable and mixed-initiative menus were more confident than those who used the static and adaptive ones. In addition, the majority of subjects (nine) who worked with the adaptive menu appeared worried and confused. During interviews after the experiment, they said that moving items made them uncomfortable and confused, so that they spent time on understanding what was happening around them. As for subjects who participated in the evaluation of the static condition, they became bored because it took them a long time to complete their tasks. In addition, it was apparent that subjects spent less time in customisation in the mixed-initiative than the adaptable condition.

3.7.1 Adaptable platform

All 15 participants customised the adaptable menus based on the instructions provided, but most needed some encouragement and most did not appear to wish to customise all systems. The average time spent on customisation (N=15) was 24 minutes. Given the choice on the home page between a full-featured interface with suggested content and a reduced interface where they would add their own content, 12 participants chose the full-featured interface and the remaining 3 chose the vacant one. No participant changed from one type of keyboard to another, all choosing to use the default QWERTY keyboard. Furthermore, no participant changed the buttons on the keyboard. After the experiment, participants stated that they appreciated the customisations and had started to realise the benefit of customising the system.

3.7.2 Adaptive platform

All the 15 participants in the adaptive condition registered with the system according to the instructions provided. When offered a choice of keyboards, 13 chose the QWERTY and 2 the alphabetical format, while none selected the AZERTY format or QWERTY with keypad.

3.7.3 Mixed-Initiative platform

In the mixed-initiative condition, participants appeared to like to customise the number of items displayed. All fifteen registered with the system according to the instructions. Keyboard preferences were similar to those for the adaptive condition: 14 for QWERTY and one for alphabetical. The average time spent on customising was 6 minutes (N=15).

3.7.4 Efficiency

Efficiency was measured in terms of task accomplishment time. Figures 11 and 12 compare the four environments in terms of efficiency on the main tasks, consisting of two types: searching for items and purchasing by keyboard. The significance of differences in task accomplishment time was tested using one-way ANOVA, which revealed significant differences in efficiency among the four groups in searching tasks ($F(3, 56) = 26.7$, $P < .001$). In addition, a t-test was carried out in order to ascertain if there was a significant difference between the conditions.

Figure 11 shows the mean value of time taken to search for the required items in the main tasks. It can be seen that subjects under the static condition took longest to search for items at all complexity levels. In particular, in complex tasks these subjects spent approximately double the time required by those working under the other conditions, while in the medium-level tasks subjects using the adaptable approach spent approximately twice as long as those using the adaptive and mixed-initiative approaches, both of which were composed of personalised items which required less time to process. In the adaptive condition, subjects spent the same amount of time on easy as on medium-level tasks, probably because they first had to learn how the adaptive menu worked, and when they had become familiar with it on the easy task they could then perform faster than they otherwise would have at the next level. Finally, subjects in the mixed-initiative condition performed fastest overall, since they had the benefit of the adaptive feature (recently clicked items) but with less confusion, as they could control undesirable modifications. The significance of the differences was tested using the t-test. Table 7 shows that significant differences were found between mixed-initiative and other conditions. By contrast, there was no significant difference between the adaptable and adaptive conditions (t25=1.4, cv =2.06).

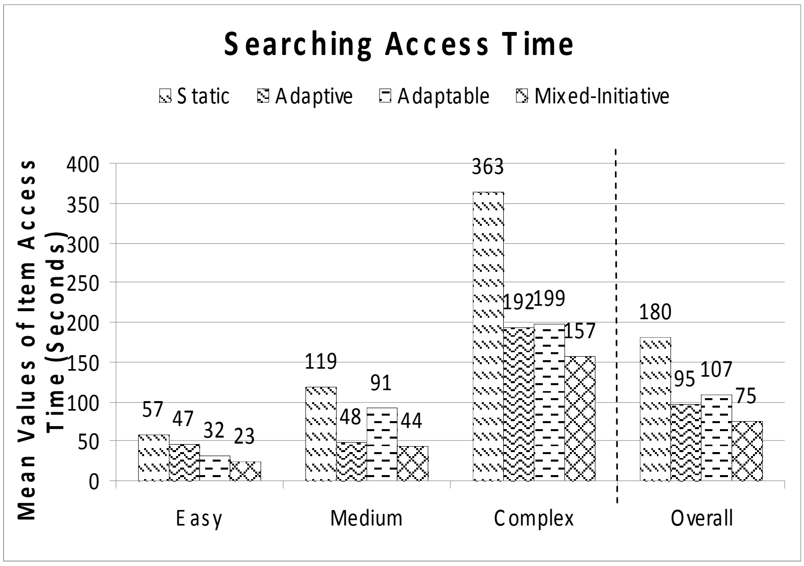


Figure 11: Efficiency in searching (main tasks) in terms of mean accomplishment time

Figure 12 shows the mean value of time taken to purchase the required items by keyboard. It can be seen that for complex tasks, subjects took longer to purchase items under the adaptive condition than any other conditions. A one-way ANOVA revealed significant differences in efficiency among the four groups in the purchasing tasks ($F(3, 56) = 23.2$, $P < .001$). In addition, the t-test was used to compare the four experimental conditions.

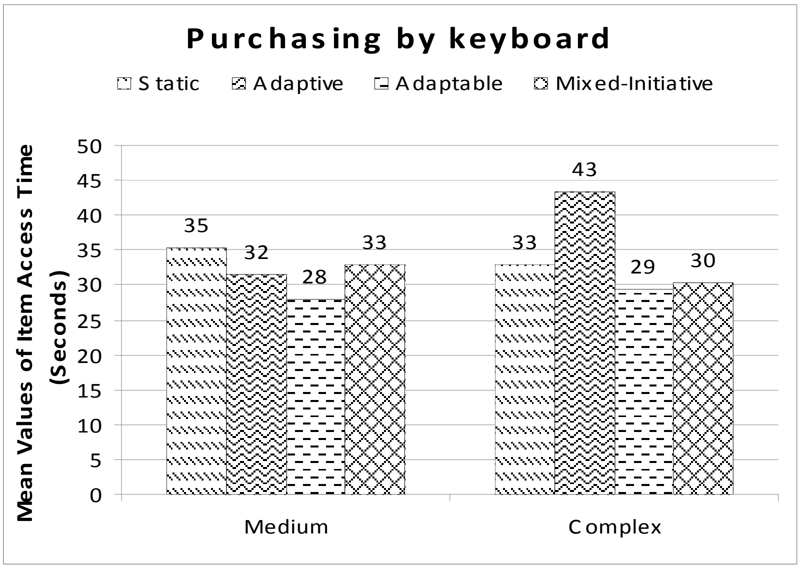


Figure 12: Efficiency in keyboard (main tasks) in terms of mean accomplishment time

Table 7: Results of t-test for main tasks (statistically significant results in bold)

Conditions	Searching	Purchasing
Adaptable vs. Adaptive	t25=1.4, cv=2.06, r = 0.27	t23=1.5, cv=2.07, r = 0.3
Adaptive vs. Static	**t17 =5.2, cv=2.11, r = 0.78**	**t16=5.1, cv=2.12, r = 0.79**
Adaptive vs. Mixed-Initiative	**t24 = 3.6, cv = 2.06, r = 0.59**	**t28=2.8, cv=2.04, r = 0.47**
Adaptable vs. Static	**t19 = 4.3, cv = 2.09, r=0.70**	**t20=4.07, cv=2.08, r = 0.67**
Adaptable vs. Mixed-Initiative	**t19 = 4.3, cv = 2.09, r = 0.70**	**t22=3.6, cv=2.07, r = 0.61**
Mixed-Initiative vs. Static	**t15 = 6.7, cv = 2.1, r = 0.87**	**t16=6.1, cv=2.12, r = 0.84**

As Table 7 shows, there was a significant difference between the conditions in purchasing tasks. Subjects who purchased items by keyboard under the adaptable condition were fastest, closely followed by the mixed-initiative and then by the static condition. Surprisingly, subjects took longer to purchase items under the adaptive than all other conditions, since the complex tasks consisted of two impersonal items.

Figures 13 and 14 show differences between the four environments on the training tasks. In regard to task accomplishment time, the significance of the differences was tested using one-way ANOVA, which revealed significant differences in efficiency between the four groups in searching and purchasing by keyboard ($F(3, 56) = 13.3$, $P < .001$; $F(3, 56) = 4.9$, $P < .004$ respectively). In addition, a t-test was carried out in order to ascertain if there were significant differences between the conditions.

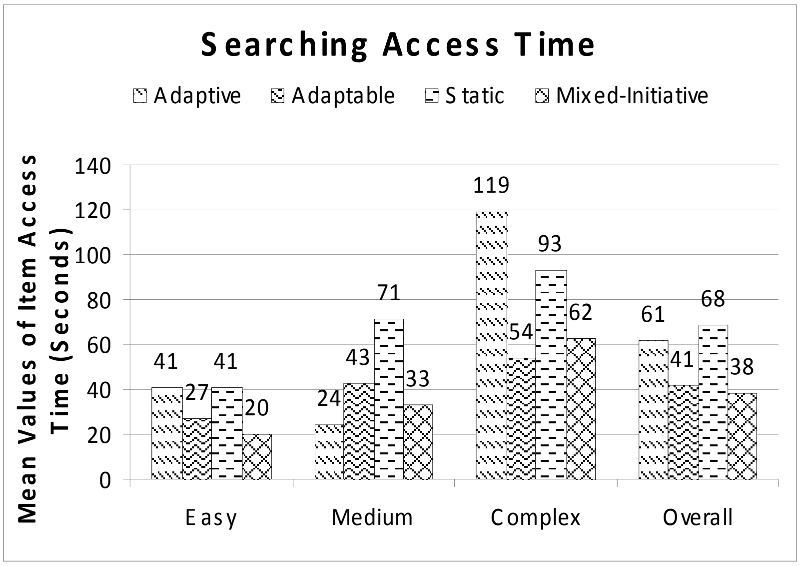


Figure 13: Efficiency in searching (learning tasks) in terms of mean accomplishment time

Figure 13 shows that on the complex tasks, subjects took more or less double the time under the adaptive approach as under either the adaptable or mixed-initiative approaches. On the other hand, on medium-level tasks, subjects spent three times as long under the static condition as the adaptive one, and twice as long as under the mixed-initiative condition. Overall, the mixed-initiative and adaptable approaches were more efficient than the adaptive and static ones.

Figure 14 shows that on medium-level tasks, subjects spent more than double the time under adaptive conditions as under the mixed-initiative approach. Overall, the most efficient approach was mixed-initiative, followed by adaptable.

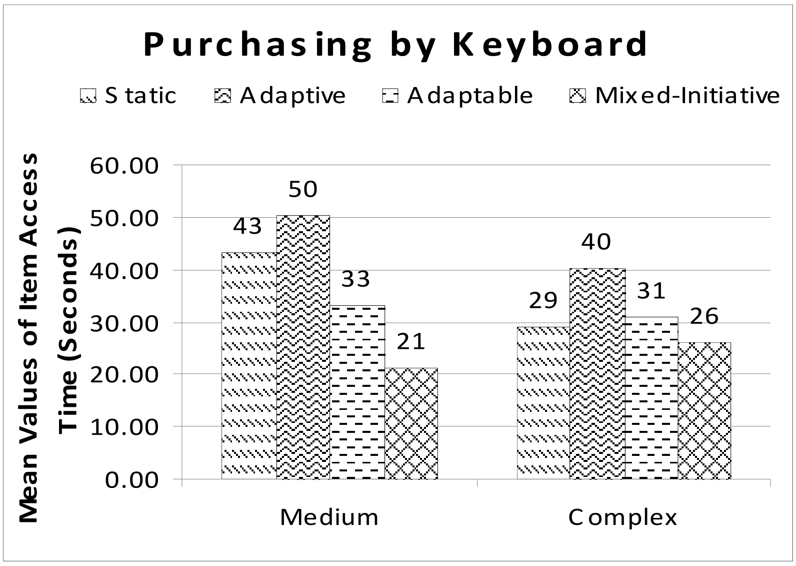


Figure 14: Efficiency purchasing in (learning tasks) in terms of mean accomplishment time

Table 8: T-test results for training tasks (statistically significant results in bold)

Conditions	Searching	Purchasing
Adaptable vs. Adaptive	**t20=3.7, cv=2.08, r = 0.64**	t15 =1.8, cv=2.1, r = 0.42
Adaptive vs. Static	t28 =0.95, cv=2.04, r = 0.18	t21 =1.1, cv=2.08, r = 0.23
Adaptive vs. Mixed-Initiative	**t21 = 4.2, cv = 2.08, r = 0.68**	**t16 = 3.04, cv = 2.12, r = 0.61**
Adaptable vs. Static	**t19 = 4.5, cv = 2.09, r = 0.72**	t19 =1.08, cv=2.09, r = 0.24
Adaptable vs. Mixed-Initiative	t27 = 0.84, cv =2.05, r = 0.16	**t27= 3.66, cv = 2.05, r = 0.58**
Mixed-Initiative vs. Static	**t20 = 4.9, cv = 2.08, r = 0.74**	**t21= 3.2, cv = 2.08, r = 0.57**

Table 8 indicates that there were significant differences between the mixed-initiative and both the adaptive and static conditions in searching and purchasing by keyboard. Furthermore, there was a significant difference between mixed-initiative and adaptable solely in keyboard purchasing tasks, and significant differences between the adaptable and both adaptive and static conditions, solely in searching tasks.

On overall performance, subjects were faster in the mixed-initiative condition, spending an average total of 3 minutes 43 seconds, followed by the adaptable and adaptive conditions (4 minutes 29 seconds and 5 minutes 14 seconds respectively; see Table 9). Subjects were slower in the static condition, spending an average of 6 minutes 28 seconds. The difference between the conditions was significant. The results were no different when considering keyboard performance only. Subjects spent less time purchasing by keyboard in the mixed-initiative condition (1 minute 50 seconds) than adaptable (2 minutes 1 second), adaptive (2 minutes 36 seconds) and static (2 minutes 20 seconds). Furthermore, the results were similar when considering task performance only: subjects spent an average total of 1 minute 53 seconds in the mixed-initiative condition, followed by adaptable, adaptive and static respectively. Table 10 shows a summary of the average time and standard deviation for each condition. ST = Static, AD = Adaptable, AV = Adaptive, MI = Mixed Initiative.

Table 9: Overall Performance

Dependent variable	Mean				Standard Deviation			
	ST	AD	AV	MI	ST	AD	AV	MI
Task Performance (minutes)	4:08	2:28	2:36	1.53	31.42	26.07	18.71	12.07
Keyboard Performance (minutes)	2:20	2:01	2:36	1:50	7.12	7.63	10.68	6.82
Overall Performance (minutes)	6:28	4:29	5:14	3:43	38.54	33:7	29.39	18.89
Customisation Time (minutes)	-	25:42	-	6.58	-	2:36	-	7:18

Table 10: Mean number of clicks and pages visited

	Conditions	Static	Adaptive	Adaptable	Mixed-Initiative
Learnable Tasks	Mouse Clicks	6	4	4	4
	Visited Pages	5	3	3	2
Main Tasks	Mouse Clicks	59	21	31	26
	Visited Pages	48	13	20	19
Keyboard	Mouse Clicks	124	35	33	28

Table 10 shows that for the searching tasks, the highest average number of mouse clicks per task (59) and pages visited per task (48) were in the static condition, followed by the adaptable environment, with 31 clicks and 20 pages per task, then mixed-initiative (26 clicks and 19 pages) and adaptive (21 clicks and 13 pages per task), since fewer functions were available in the adaptive environment than in the others. It was also noticed during the experiment that subjects spent more time without moving the mouse, while they tried to understand the modifications made by the system.

In other words, the highest average number of mouse clicks (6) and pages visited (5) per task was in the static environment, followed by the adaptive and adaptable conditions, with 4 clicks and 3 pages per task in each case. The lowest number of mouse clicks and pages visited was in the mixed-initiative approach (4 and 2 per task). In the keyboard tasks, the highest average number of mouse clicks per task (124) was in the static environment, followed by the adaptive (35) and adaptable (33) conditions. The lowest average number of mouse clicks per task (28) was in the mixed-initiative condition.

3.7.5 Errors

Subjects in the mixed-initiative condition made fewer errors than those in the other conditions. The highest number of errors was made in the static condition, followed by the adaptive, adaptable and mixed-initiative conditions (see Table 11). Errors occurred in the static condition principally because subjects selected incorrect items, having spent a long time searching for items because there was no categorisation. Similarly, errors occurred in the adaptive condition principally because subjects felt lost as a result of the adaptation characteristics.

Table 11: Frequency of user errors under the different conditions

Conditions	Static	Adaptive	Adaptable	Mixed-Initiative
Selecting incorrect item	56	10	7	3
Incorrect path	14	5	3	2
Confusion	9	13	3	2
Number of failed tasks	5	3	0	0
Selecting incorrect category	4	2	3	1
Keyboard typing error	4	2	3	4
Total	92	35	19	12

3.7.6 Effectiveness

More subjects completed easy, medium and complex tasks using the mixed-initiative approach than under any other conditions, excluding medium tasks under the adaptive condition. The t-test was used to ascertain overall differences between the four conditions in number of subjects who completed all tasks. It showed that there was a significant difference at a level of 0.05 between the mixed-initiative and both adaptable ($t(3)=4.3$, $cv=3.1$, $r = 0.93$) and static ($t(3)=12.3$, $cv=3.1$, $r = 0.99$) conditions, but an insignificant difference between mixed and adaptive ($t(3)=2.04$, $cv=3.1$, $r = 0.76$). More subjects completed easy, medium and complex tasks using the adaptable and adaptive environments than the static one (see Figure 15). Completion of easy tasks was also found to be higher for the adaptable than the adaptive condition, but lower for medium tasks and similar for complex tasks.

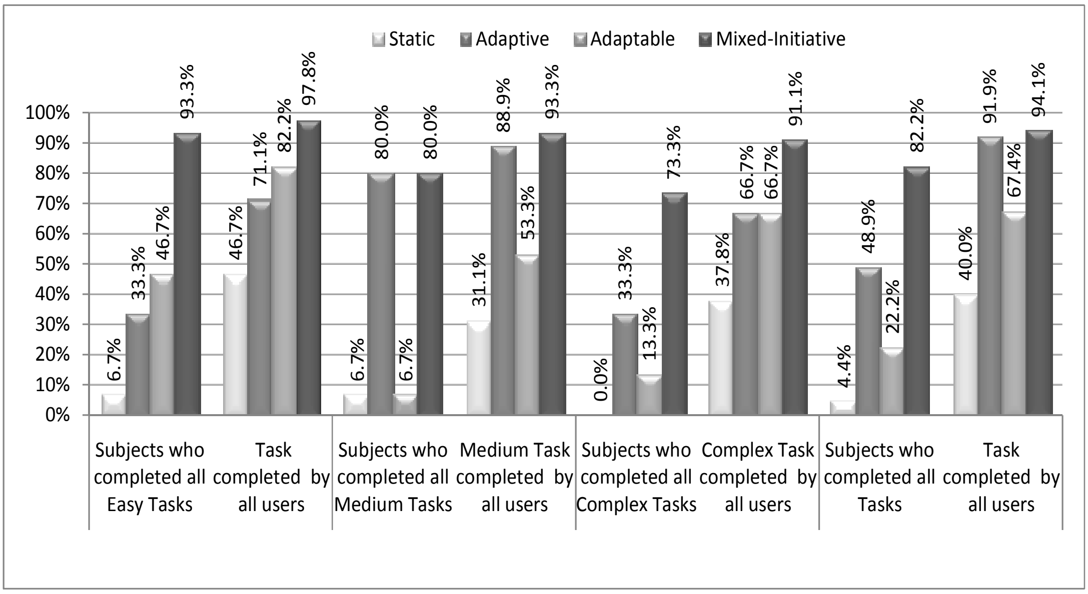


Figure 15: Effectiveness in Searching (main tasks)

Overall, there was a significant difference in the numbers of subjects who completed all tasks between the adaptable and static conditions (t (2) = 1.4, cv = 4.3, r = 0.70) and between adaptive and static (t (2) = 2.8, cv = 4.3, r = 0.91), but no significant difference in the number of subjects who completed all tasks between the adaptable and adaptive environments (t (4) = 1.34, cv = 2.7, r = 0.56).

Figure 15 shows the percentages of easy, medium and complex tasks completed by all subjects in each of the four conditions, as well as the overall percentages. The numbers of tasks not completed were as follows: 8 tasks (1 easy, 3 medium, 4 complex) under the mixed-initiative condition, 33 tasks (13 easy, 5 medium, 15 complex) under the adaptive, 44 tasks (8 easy, 21 medium, 15 complex) under the adaptable and 83 tasks (24 easy, 31 medium, 38 complex) under the static condition. A t-test was used to determine the diversity between the four conditions. Table 12 shows t-test results indicating a significant difference at 0.05 between the number of tasks completed by all subjects using the mixed-initiative compared to static (t(3) = 11.3, cv = 3.1, r = 0.99) but not to adaptable (t(2) = 2.6, cv = 4.3, r = 0.88) or adaptive (t(2)=3.1, cv = 4.3, r = 0.91). There was also a significant difference in the numbers of tasks completed by all subjects between the adaptive and static conditions (t(4) = 4.5, cv = 2.7, r = 0.91), but no statistically significant difference between static and adaptable (t (3) = 3.04, cv=3.1, r = 0.87) or between adaptable and adaptive (t (4) = 0.757, cv = 2.7, r = 0.35).

Table 12: T-test results for completion (statistically significant ones in bold)

Conditions	Subjects who completed all tasks	Tasks completed by all subjects
Adaptable vs. Adaptive	t4=1.34, cv=2.7, r = 0.56	t4 = 0.757, cv = 2.7, r = 0.35
Adaptive vs. Static	t2 =2.8, cv=4.3, r = 0.89	**t4 = 4.5, cv=2.7, r = 0.91**
Adaptive vs. Mixed-Initiative	t3 = 2.04, cv = 3.1, r = 0.76	t2=3.1, cv=4.3, r = 0.91
Adaptable vs. Static	t2 =1.4, cv=4.3, r = 0.70	t3 = 3.04, cv=3.1, r = 0.87
Adaptable vs. Mixed-Initiative	**t3 = 4.3, cv = 3.1, r = 0.93**	t2 = 2.6, cv = 4.3, r = 0.88
Mixed-Initiative vs. Static	**t3 = 12.3, cv = 3.1, r = 0.99**	**t3=11.3, cv=3.1, r = 0.99**

Overall, just 8 subjects did not complete all tasks under the mixed-initiative condition, whereas 23 subjects failed to do so when using the adaptive approach, as did 24 in the adaptable condition. Under the static environment, by contrast, only 2 subjects did complete all tasks. This shows that the overall number of subjects who completed all tasks was higher in the mixed-initiative than the other conditions. Fewer completed all levels of task using static menus than other types. Table 12 shows a significant difference in the number of subjects who completed the tasks at a significance level of 0.05 ($F = (3, 11)$, $p <0.004$). To compare the effectiveness of learning tasks between the four conditions, the critical time for learning task completion was derived. Thus, a task would be regarded as successfully completed if subjects finished it within this time. Figure 16 shows differences between the four conditions in the number of subjects who completed learning tasks and the number of learning tasks completed by all subjects. This indicates that the effectiveness of training tasks varied among the four conditions. The difference was found to be statistically significant at 0.05 using ANOVA.

Eleven subjects completed all training tasks under the adaptable condition, followed by mixed-initiative (9 subjects), static (3) and adaptive (2). The main reason for the poor results of the adaptive condition is that the position of items in all lists was subject to persistent change, which caused confusion among subjects. Figure 16 shows that the percentage of subjects who completed all tasks using the mixed-initiative condition was higher than the adaptive and static conditions but not higher than the adaptable condition. In addition, Figure 16 shows that over 90 % of training

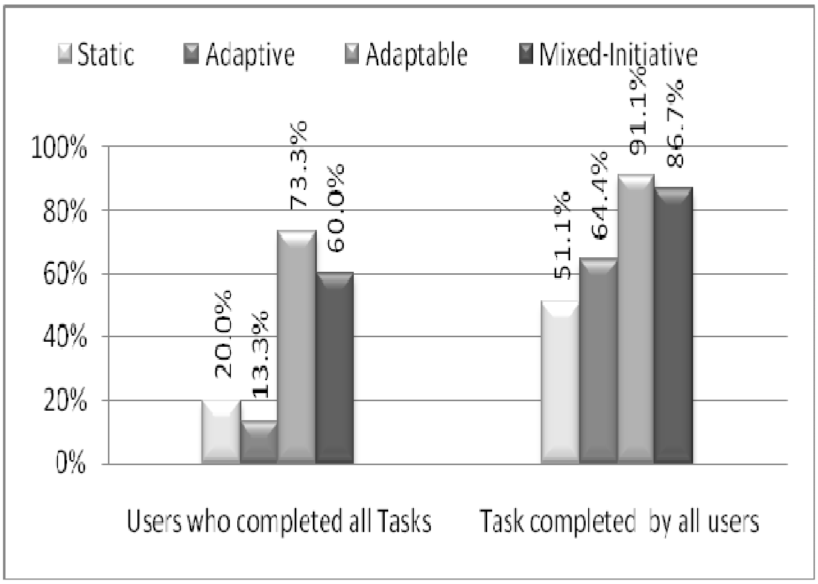


Figure 16: Effectiveness in searching (training tasks)

tasks were completed by all subjects using the adaptable menus, almost as many under the mixed-initiative condition, around two-thirds under the adaptive environment and half in the static condition.

3.7.7 Satisfaction

A questionnaire with 12 statements and a 6-point Likert rating scale was used to elicit subjects' views after performing each experimental condition. Subjects had to select a response from 1 to 6, where 1 indicated strong agreement and 6 strong disagreements with measures of overall satisfaction under each condition. Table 13 shows the data obtained along with the frequency and the percentage of subjects who agreed and disagreed with each statement in the four conditions. Subjects who selected responses 1 to 3 were considered to agree and the others to disagree. The table shows a mix of positive and negative statements. The former were on ease of use, ease of search, ease of purchase, ease of navigation, ease of control, suitability and overall satisfaction, while the negative ones were that the system was difficult to learn, confusing, made the user feel nervous and had a deficiency of control.

The mean value of subjects' responses regarding ease of searching inside the item lists, ease of navigation, ease of shopping, ease of control and overall satisfaction (statements 3, 4, 5, 6 and 12) was higher for the mixed-initiative than other conditions. Table 13 shows that the mean value of subjects' responses regarding ease of use and ease of shopping (statements 1 and 5) was higher for adaptable and

Table 13: Subjects' agreement and disagreement with each statement for the four conditions (frequency and percentage)

Statements	Static		Adaptable		Adaptive		Mixed-Initiative	
	Disagree	Agree	Disagree	Agree	Disagree	Agree	Disagree	Agree
Easy to use	9 (20.0%)	36 (80.0)%	2 (4.4%)	43 (95.6%)	8 (26.7%)	37 (73.3%)	2 (4.4%)	43 (95.6%)
Easy to purchase	7 (15.6%)	38 (84.4%)	4 (8.9%)	41 (91.1%)	9 (28.9%)	36 (71.1%)	4 (8.9%)	41 (91.1%)
Ease to search inside the lists	16 (35.6%)	29 (64.4%)	3 (6.7%)	42 (93.3%)	11 (40.0%)	35 (60.0%)	2 (4.4%)	43 (95.6%)
Ease to navigate	6 (13.3%)	39 (86.7%)	4 (8.9%)	41 (91.1%)	8 (26.7%)	38 (73.3%)	0 (0.0%)	45 (100%)
Ease to shop	13 (28.9%)	32 (71.1%)	2 (4.4%)	43 (95.6%)	7 (42.2%)	37 (57.8%)	2 (4.4%)	43 (95.6%)
Easy to control	10 (22.2%)	35 (77.8%)	1 (2.2%)	44 (97.8%)	7 (24.4%)	37 (75.6%)	0 (0.0%)	45 (100%)
Difficult to learn	23 (51.1%)	22 (48.9%)	38 (84.4%)	7 (15.6%)	40 (77.8%)	5 (22.2%)	45 (100%)	0 (0.0%)
Feel nervous	15 (33.3%)	30 (66.7%)	37 (82.2%)	8 (17.8%)	34 (57.8%)	11 (42.2%)	39 (86.7%)	6 (13.3%)
Feel confusing	17 (37.8%)	28 (62.2%)	38 (84.4%)	7 (15.6%)	33 (55.6%)	12 (44.4%)	42 (93.3%)	3 (6.7%)
Suitability	17 (37.8%)	28 (62.2%)	38 (84.4%)	7 (15.6%)	31 (55.6%)	14 (44.4%)	33 (73.3%)	12 (26.7%)
Deficiency of control	5 (11.1%)	40 (88.9%)	34 (75.6%)	11 (24.4%)	23 (28.9%)	22 (71.1%)	41 (91.1%)	4 (8.9%)
Satisfaction	12 (26.7%)	33 (73.3%)	2 (4.4%)	43 (95.6%)	9 (33.3%)	37 (66.7%)	0 (0.0%)	45 (100%)

mixed-initiative conditions than adaptive and static ones. As for the negative statements (7, 8, 9 and 11), only 13.3% of mixed-initiative subjects felt nervous, whereas 6.7% felt confused and all subjects found the platform easy to learn. In the adaptable condition, 17.8% of subjects felt nervous, while 15.6% felt confused and 15.6% found the platform difficult to learn. In the adaptive condition, 51.1% of subjects felt nervous, 37.7% felt confused and 60% found the platform easy to learn. In the static condition, 66.7% of subjects felt nervous, 62.2% felt confused and 48.9% found the platform difficult to learn.

3.7.8 Controllability

At the end of each session subjects were asked to give ratings of 1 to 10 on a Likert scale for user control and for website control. Figure 17 shows the differences among the four conditions. The highest scores for subject control were around 90% for the mixed-initiative and adaptable conditions. On the other hand, in terms of website

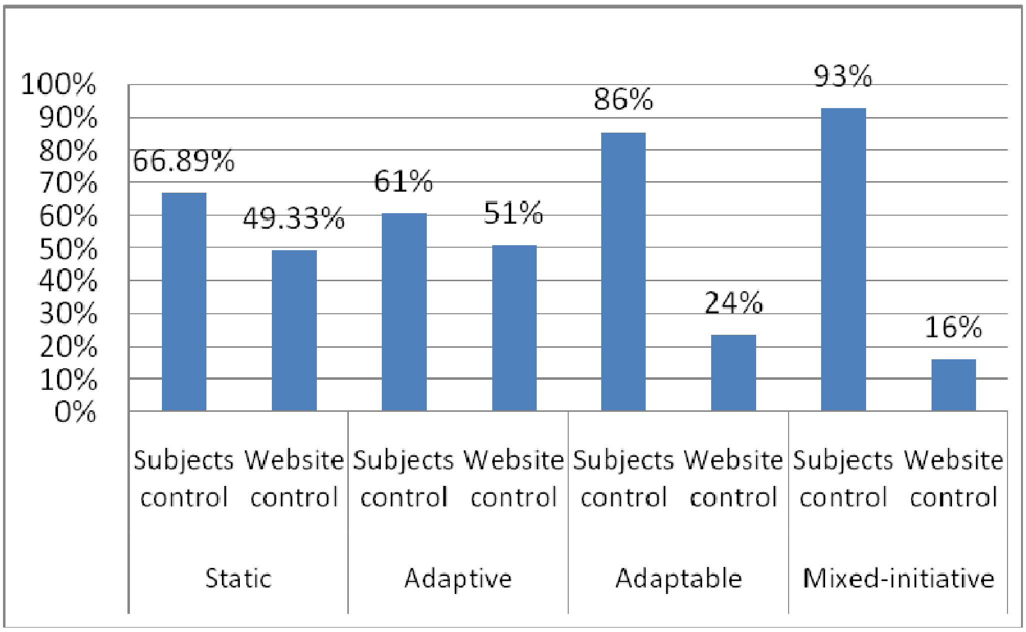


Figure 17: Controllability

control mixed-initiative had the least score. Closely followed by adaptable condition. However, there was a slight difference between subjects control and website control. Subjects who utilised the mixed-initiative had more control on their condition than other one. Followed by the adaptable condition, static, and adaptive with (86%), (66.89%), and (61%) respectively.

3.7.9 Customisation Behaviour

Subjects had two opportunities to customise: one before starting to use the platform and the other before performing each task. In addition, they were allowed to customise during the tasks, if they had forgotten to do so, but this time was included in the task time. According to Figure 18 subjects who customise the adaptable condition spent four times more minutes than those who customise the mixed-initiative condition. In other words, subjects spent significantly less time customising the mixed-initiative platform than the adaptable platform with averages of 6 minutes 58 seconds and 25 minutes 42 seconds respectively. t-Test results showed that there was a significant difference at 0.05 between the time spend to customise the adaptable and mixed-initiative conditions ($t_{14} = 9.32$, $p < 0.05$, $r = 0.928$).

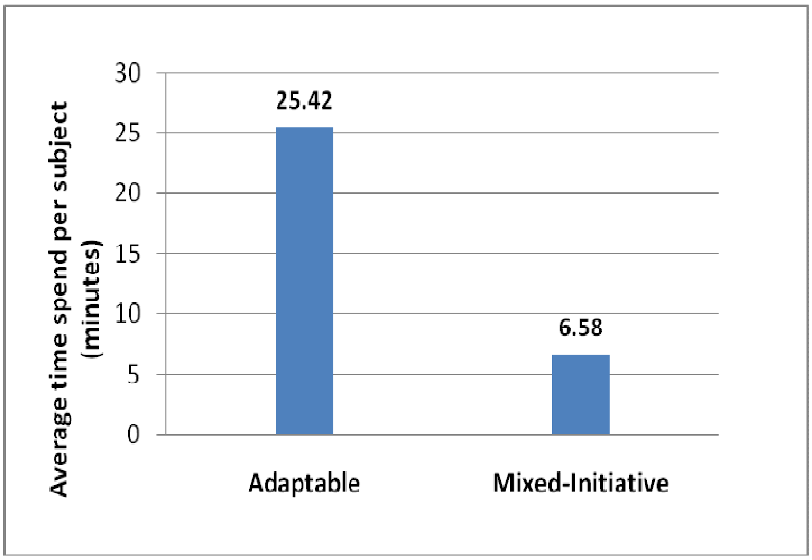


Figure 18: Customisation Time

3.8 Discussion

This chapter has documented the initial experiment carried out in order to investigate the efficiency, effectiveness and satisfaction of different types of personalisation approaches, including a static one, to web content. The main reason for this experiment was that an examination of the current research on personalisation reveals conflicting findings as to which of these approaches is best (most usable). In addition, far too little attention has been paid to comparing the adaptable, adaptive and mixed-initiative approaches. The metrics used to measure efficiency were task accomplishment time and frequency of error-occurrence, while effectiveness was measured by calculating the number of subjects completing all tasks, and the number of tasks completed successfully within task criterion times. Satisfaction was measured by using 6-point Likert scales. The results of this study provide encouraging evidence that the mixed-initiative approach more usable (more efficient, effective, and preferable) over the others conditions. This indicated that subjects preferred to have full control, at the same time as receiving some suggestions and assistance from the system. In other words, mixing the adaptive and adaptable approaches is far more helpful than just providing either of these techniques alone.

Since the adaptive, adaptable and mixed-initiative approaches have different levels of controllability, this experiment was conducted to address some questions concerning controllability. For example, how much control users actually feel whilst utilising adaptive, adaptable, and mixed-initiative approach. More specifically, is this control is enough to do their tasks easily. Therefore, we asked subjects after performing each level of tasks (easy, medium, and complex) along with at the end of the experiment. In addition, the experimental results were obtained from both quantitative and qualitative measures, along with self-reported and observed data. In addition, an interview was conducted with subjects when needed. The results indicate that providing more control than users required caused confusion and irritation. For example, in the adaptable condition, subjects had full control of their content, whereas they had less control under the mixed-initiative approach. In addition, subjects spent significantly more time customising the adaptable platform than the mixed-initiative platform with averages of 25 minutes 42 seconds and 6 minutes 58 seconds respectively. This should provide more controllability feelings on subjects who utilised the adaptable approach. In the adaptable approach, the majority of subjects (12) did not wish to customise their environment fully. In contrasts, subjects were happier to control the system in the mixed-initiative environment. In addition, the data shows that the customisation time in the mixed-initiative case was significantly lower than in the adaptable condition, although the highest scores for subject control were 93% for the mixed-initiative and 86% for the adaptable conditions. In addition, it was noticeable that subjects who participated in the evaluation of the mixed-initiative were more confident than under static, adaptable and adaptive conditions. For example, the majority of subjects (Nine) who participated in the adaptive conditions looks worried and confused. After, the experiment during the interview, they said that moving items makes them not comfortable. This confusion made them spending time on comprehension what is happening around them. Furthermore, subjects who participated in the evaluation of the static condition get bored because they spending long time to complete their tasks. In addition, it was apparently noticeable that subjects spent less time in customisation in the mixed-initiative than the adaptable conditions.

This experiment aimed to discover which personalised approach users preferred. In addition, we were interested in the impact of personalisation approaches on the e-commerce environment. Therefore, we examined the ease of searching inside the lists of items, ease of navigation, ease of shopping, ease of control, learnability, nervousness, confusion, suitability and overall satisfaction (see Table 13). The results indicate that user satisfaction was affected by the level of controllability provided to users, along with the effort required from users to do their tasks. For example, subjects' responses regarding ease of searching, ease of navigation, ease of shopping, ease of control and overall satisfaction (statements 3, 4, 5, 6 and 12 in Table 13) were higher for the mixed-initiative than other conditions. As for the negative statements (7, 8, 9 and 11 in Table 13), scores were lower for the mixed-initiative than other conditions. This means that providing more or less control than what users expected and needed reduced satisfaction levels. In addition, the results show that subjects did their tasks with less effort in the mixed-initiative condition than any other condition. More specifically, just 8 subjects did not complete all tasks under the mixed-initiative condition, whereas 23 subjects failed to do so when using the adaptive approach, as did 24 in the adaptable condition. Under the static environment, by contrast, only 2 subjects completed all tasks. This result (see Section 3.7.6) and the results for controllability (see Section 3.7.8) indicate that users preferred to have control, but once the level of control required too much attention and effort, users began to dislike it.

There was a variety of responses to the design of each approach. First of all, in terms of design of the adaptive interface, subjects generally liked the way that the system assisted them by moving items to the top. However, there were comments suggesting that moving items repeatedly was confusing. In other words, there was a need for adaptation but with less movement. On the other hand, in terms of design of the adaptable interface, subjects generally liked the way of controlling the number of items displayed in each list and controlling the contents by dragging and dropping items from one part to another, along with adding new contents to the main home page. Furthermore, subjects were aware of the number of items displayed in each list. However, during the interviews some commented that displaying all items in each part would make the search very difficult. In other words, the interface would become visually too complex. Last but not least, in terms of design of the mixed-initiative interface, subjects generally liked locking lists to prevent items from moving up and down, unlocking them when required. This confirmed that our solution is generally acceptable. Some subjects suggested that the device of locking items could be improved if the system provided some recommendations. Ultimately, during the experiment it was noticeable that subjects were willing to accept suggestions from the system while performing their tasks.

3.9 Summary

This chapter has described the initial comparative evaluation of static, adaptive, adaptable and mixed-initiative approaches to e-commerce environments. The results indicate that overall, there was a significant difference between personalisation approaches in each usability parameters (efficiency, effectiveness and satisfaction). These differences are critical because they can cause confusion and irritation to users. The overall results indicate that subjects preferred the mixed-initiative condition over the other three, that they performed fastest and felt that they had most control while using the mixed-initiative approach. The initial experiment has indicated many questions in need of further investigation, such as whether we could mitigate the drawbacks of each condition, which condition would be most usable and whether the usability of these approaches would differ with the GUI. Therefore, more research needs to be undertaken on this topic to explore these approaches from different perspectives.

This chapter has attempted to answer the first part of question one presented in Section 1.3. In order to fulfil the second aim of this study, the second part of that question will be discussed next. The next chapter describes and evaluates these approaches under the desktop environment. More specifically, it is concerned with testing the effects of these approaches on one of the main interface components (the menus). Examining the personalisation approaches in different domains will show any other differences between adaptive, adaptable and mixed-initiative approaches. In addition, it will provide a deeper understanding of the behaviour of these approaches.

Chapter 4: Study Two: Comparison of Personalised Approaches to Graphical User Interfaces

4.1 Introduction

This chapter seeks to address the second part of question one, stated in Chapter 1, by reporting on a second experiment to assess the usability of personalised conditions in a graphical user interface. These conditions are adaptable, adaptive split (split), adaptive/adaptable highlighted (highlighted), adaptive/adaptable minimised (minimised) and mixed-initiative menus. More specifically, it compares the usability of these five types with regard to task accomplishment time, frequency of error-occurrence, effectiveness and user satisfaction.

This chapter moves forward the debate on personalisation by addressing the question: Does the size of personalised information content affect efficiency, effectiveness, user satisfaction and customisation behaviour? The work reported here investigated empirically the effects of content size on five different personalised menu types. In order to carry out this comparative investigation, two independent experiments were conducted, on small menus (17 items) and large ones (29 items) respectively. These were tested dependently using 30 subjects each. The empirical evaluation of these two experiments is expected to provide a profound understanding of personalised conditions and the question in hand.

The chapter begins with a statement of the aims and objectives of the second experiment, followed by a description of the experimental platform. The hypotheses are then stated and the experimental design and methodology are described. Finally, this chapter presents an illustrative and descriptive analysis of the data obtained and discusses the results of the three usability parameters (efficiency, effectiveness and satisfaction) along with subjects' customisation behaviour, in order to answer the question above (see Figure 19).

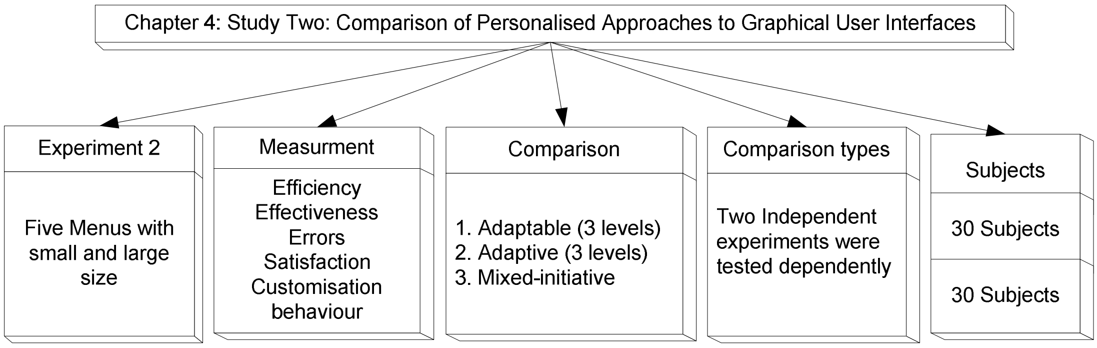


Figure 19: Summary of chapter 4

4.2 Aims

The aim of this experiment is to shine further light on the personalisation debate through an examination of the usability of graphical user interface control structure, taking menus as an example. Its second aim is to move the debate forward by proving empirically that there are some factors which cause approaches to personalisation to have positive effects at one time and negative ones at others. This will be done through an empirical comparison concerning the effect of content size on the usability of personalised conditions. More specifically, it measures the effect of small versus large menu size on the efficiency, effectiveness and satisfaction of adaptive menus (split menu and both highlighted and minimised menus in session 1), adaptable menus (adaptable and both highlighted and minimised menus in session 2) and mixed-initiative menus.

4.3 Objectives

In order to fulfil our goals, six objectives had to be attained. The first was to measure precisely the efficiency of each condition in small and large menus by timing the completion of tasks and quantifying errors under each condition. The second objective was to measure the effectiveness of each condition in small and large menus by calculating the percentage of tasks completed successfully and of subjects who successfully completed all tasks. The third objective was to measure the satisfaction of users under each condition in small and large menus by obtaining the subjects' opinions. The fourth objective was to examine subjects' customisation behaviour when using small and large menus. The fifth objective was to discover by comparing the findings of the four above objectives whether the size of content affected personalised approaches. The final objective was to investigate the effect of different levels of adaptivity (split menu and both highlighted and minimised menus in session 1) and adaptability (adaptable and both highlighted and minimised menus in session 2) on efficiency, effectiveness and satisfaction.

4.4 Experimental Approaches

Graphical user interface components such as menus, toolbars, and buttons fall under the category of control structures. We chose to investigate menus rather than other components because they are common and considered one of the main graphical user interface components. In addition, there are industrial and academic research studies of adaptive and adaptable menus, while the debate as to the relative merits of the two approaches continues. Furthermore, current research on personalisation reveals contradictory findings (see Figure 20).

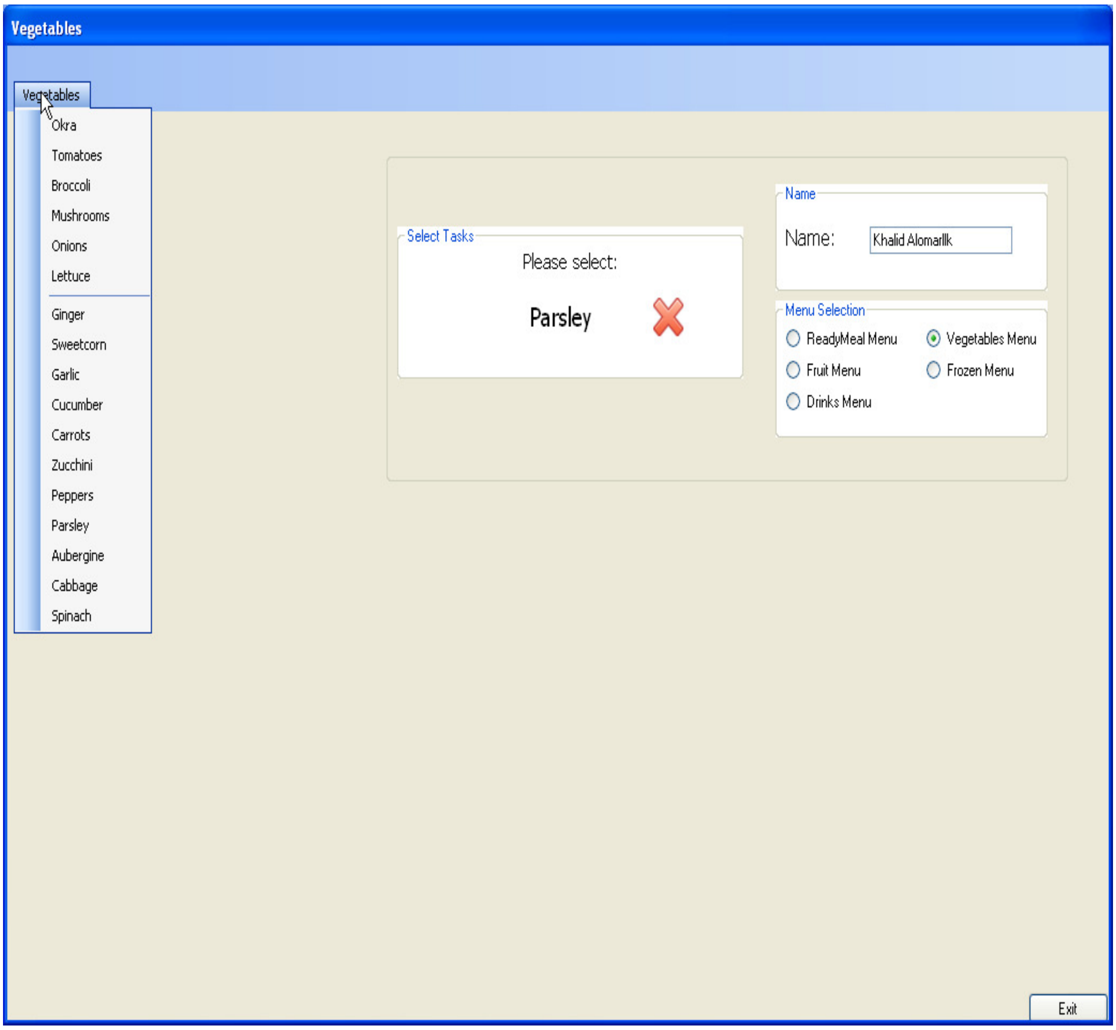


Figure 20: Screen layout in the experiment

4.4.1 Menu Design

Five different menu conditions were tested in each of two experiments (on small and large menus): adaptable, split, highlighted, minimised and mixed-initiative menus. Figure 21 illustrates the five menu types tested in experiments 1 and 2. Our work is different from others because our comparison involved a combination of different approaches. Since the division between personalised approaches is not straightforward, a mixture of these is included in the comparison (see Table 14).

The aim was to understand subjects' behaviour under the adaptive, adaptable and mixed-initiative conditions and how it varied with menu size; in other words, to explore the impact of size on these five menu conditions. Within the adaptive approach, the chosen techniques were split, highlighted and minimised menus, because their use is commonly reported in the literature with successful results and they are commercially utilised. These three techniques provided three levels of adaptation occurring mainly in session 1: (1) changes occurring without moving items (that is, highlighted menu), (2) changes made by moving recently and frequently clicked items to the top of the list and leaving the others unchanged (that is, split menu) and (3) changes made by moving only frequently clicked items to the top of the list and leaving the others unchanged (that is, minimised menu). It was considered essential to investigate which of these techniques was more usable on small and large menus. On the other hand, within the adaptable approach, the chosen techniques were (1) customisation with help not provided (that is, adaptable menu), (2) customisation with assistance provided by highlighting the frequently clicked items (that is, highlighted menu) and (3) recommendation provided by moving frequently clicked items to the top of the list, followed by a horizontal line separating the recently clicked items and hiding the others (that is, minimised menu).

Table 14: Approaches utilised in each session

Menu	Session 1	Session 2
	Approach	
Highlighted	Adaptive	Adaptable
Adaptable	Traditional	Adaptable
Minimised	Adaptive	Adaptable
Mixed-initiative	Mixed-initiative	
Split	Adaptive	

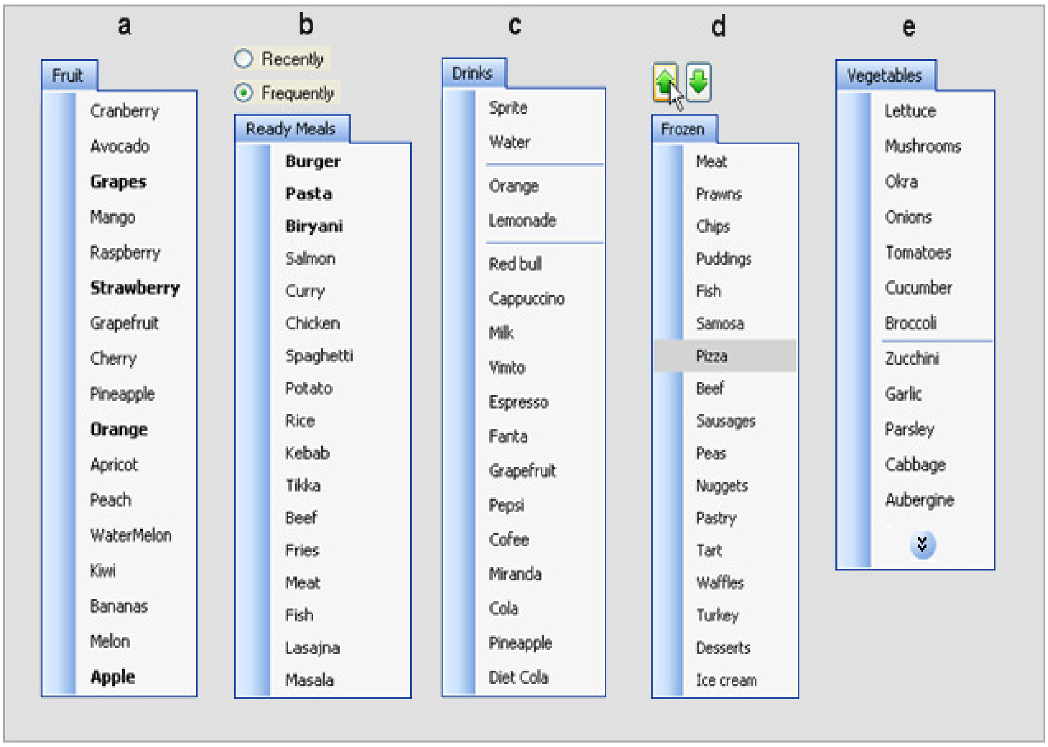


Figure 21: (a) highlighted menu; (b) mixed-initiative menu; (c) split menu; (d) adaptable menu; and (e) minimised menu.

In the adaptable condition, subjects could modify the order of items by moving them up or down. This occurred after the first session of the experiment (50 selections). The split menu was divided by two horizontal lines into three sections. The top section comprised the two most frequently selected items, the second section the two most recently selected items and the bottom section the others. The menu software counted how many times each item had been used in the 50 most recent selections and updated the list after each selection. In the highlighted menu, the most frequently selected items were boldfaced, while the others were not. After the first 50-selection session of the experiment, subjects could modify the order of items by moving them up or down. In the minimised menu, the software counted how many times each item had been used, moving frequently selected items to the top of the list and separating them from other items by a horizontal line. The top section was extendable and kept the most frequently selected items separate from the bottom section. When the user wanted to modify and customise the menu, it would be divided by two horizontal lines into three sections: the top one held the two most frequently selected items, the second comprised the two most recently selected items, while the bottom section contained the others and was hidden. Users could view the hidden items by clicking a small arrow at the end of the menu. In the mixed-initiative menu, the technique was to display the recently or frequently used items to subjects at the appropriate time. The recently selected items were displayed at the top of the menu when this feature was selected by clicking on a button labelled ‘Recently’, while the frequently selected items were displayed when the ‘Frequently’ button was clicked. Both techniques boldfaced the recently or frequently selected items and moved them to the top. Subjects were able to choose only one technique at a time but could switch from one to another at any time during the experiment. It was the subjects’ responsibility to choose the appropriate technique.

4.4.2 Menu Size

The large menu was a full-length one displayed on a large screen (see Figure 22). It contained 29 items, of which 14 were included in the experiment tasks, while the small menu contained 17 items, of which 14 were included in the experiment tasks. The small menu was the size of many menus that are commonly used and was the minimum length that would allow the same number of items (14) to be included in the tasks as for the large one, since the tasks needed to be the same on small and large menus. If there were different results between this menu size and large menus, it was expected that smaller ones would give better results.

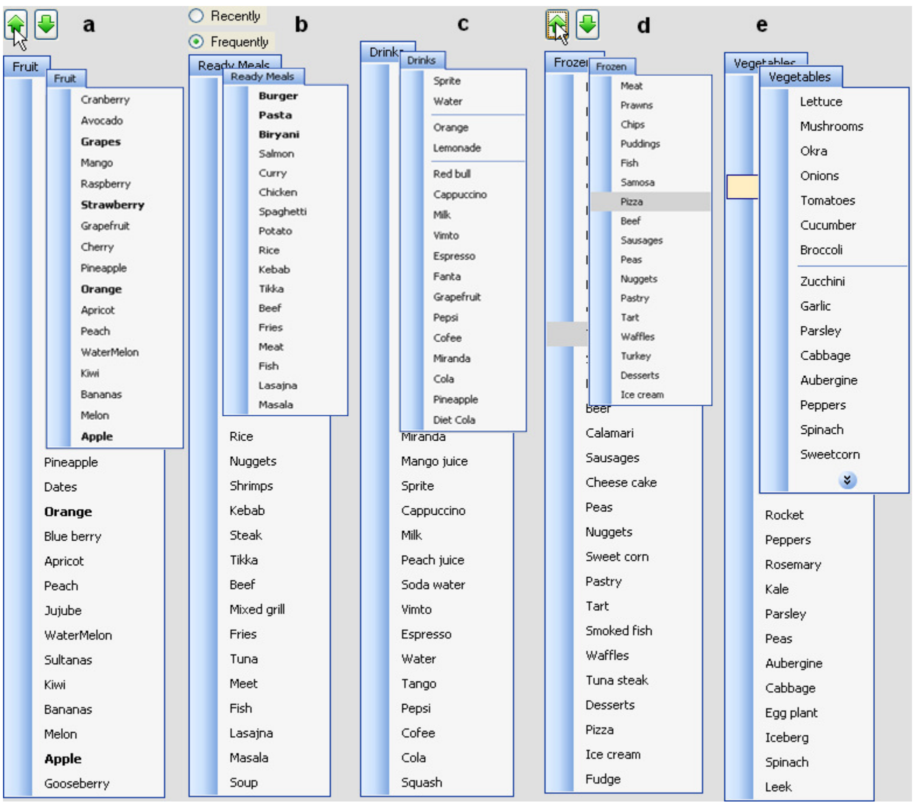


Figure 22: Menu sizes

4.5 Experimental Hypotheses

The aim of the second experiment is to measure the efficiency, effectiveness and satisfaction of the adaptive, adaptable and mixed-initiative approaches. Based on the literature review of related work, the experiment was therefore designed to test the following hypotheses.

H5: In small menus, the adaptable approach will be more efficient than the adaptive one in terms of task accomplishment time, frequency of errors, number of tasks completed successfully, and user satisfaction.

H6: In large menus, the adaptive approach will be more efficient than the adaptable one in terms of task accomplishment time, frequency of errors, number of tasks completed successfully, and user satisfaction.

H7: In small and large menus, the mixed-initiative approach will be more efficient than both the adaptive and adaptable conditions in terms of task accomplishment time, frequency of errors, number of tasks completed successfully, and user satisfaction.

4.6 Experimental Design

Each of the two experiments followed a within-subjects design and was planned to fit into a one-hour session.

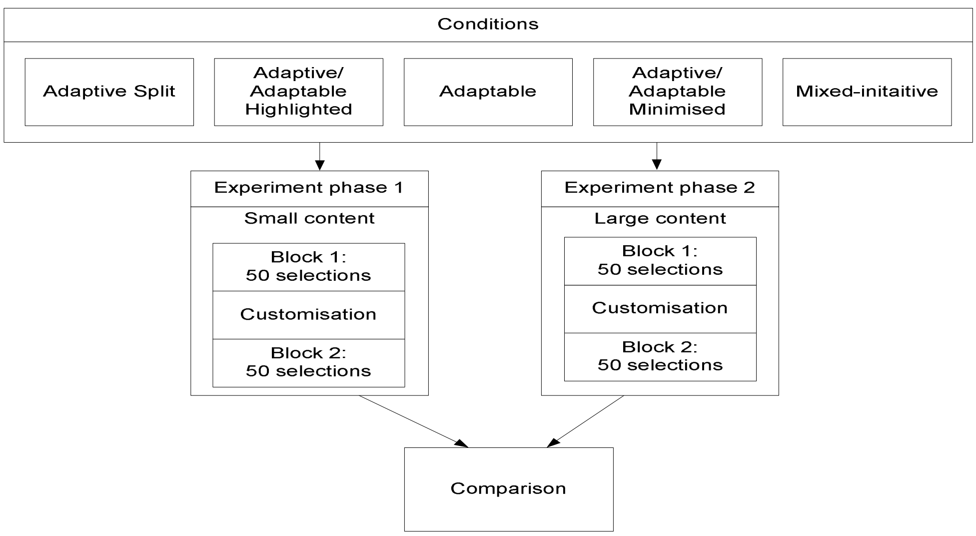


Figure 23: Plan of experiment two

Subjects were informed that the menu conditions were divided into two sessions, where session 1 consisted of a 50-item sequence selection and session 2 consisted of the identical 50-item sequence to session 1. Between the two sessions, subjects were given a 2-minute break. For the adaptable condition, subjects were allowed to take extra time during the break to customise their menus if they wished to do so. This was their only opportunity to customise (see Figure 23).

4.6.1 Evaluation Design

A two-factor mixed design was utilised: menu size (small vs. large) was tested between subjects (that is, each subject participated in one experiment), while menu type (adaptable, split, highlighted, minimised and mixed-initiative) was compared within subjects (that is, all subjects used all menus). In the first case, we chose a between-subjects design because it was essential to avoid the learning effect of using the same subject twice. In the second, by contrast, a within-subject design was preferred because the perspective of each subject in each condition was needed.

4.6.2 Subjects

A total of 60 graduate and undergraduate students voluntarily participated, 30 each on small and large menu designs. These were split 16 / 14 and 19 / 11 respectively between males and females. We decided to have 30 subjects in each experiment because we felt that this number would provide us with sufficient data on the benefits and drawbacks of each approach, while keeping the experiment under control. The ages of subjects in both experiments ranged from 18 to 44, while their average computer usage exceeded 12 hours per week. In both experiments, each subject was randomly assigned to one of five groups of 6 subjects, each of which followed the five experimental menu conditions in a different order. Subjects were given one recorded tutorial according to the experiment they participated in and were then asked to perform the same group of tasks (50 selections for each session in each condition).

4.6.3 Apparatus

For each experiment an application program was developed using Microsoft Visual Basic.Net. Personal computers with Pentium IV 1.5 GHz processors and 17-inch monitors were used in the experiment.

4.6.4 Menu Labels

In the small menu experiment, 85 different nouns from five label categories (17 nouns in each category) were used to label the menu items, while for the large menus, there were 145 different nouns from the five label categories (29 in each category). The categories in both cases were vegetables, fruits, drinks, frozen food and ready meals. Nouns shorter than four or longer than eleven characters were excluded, while no more than four nouns in any category had the same initial letter. The category name was shown in the title bar at the top of the menu.

4.6.5 Tasks

All subjects were asked to make the same number of selections (50 selections each). Each condition comprised of two task sessions, each of which contained 50 selections. Therefore, each subject performed a total of 500 selections and the thirty subjects made a total of 15 000 selections in each experiment.

4.6.6 Selection Frequency

Table 15 shows the distribution of the selection frequencies used in the two experiments. The numbers in the first, fourth and seventh columns of the table indicate the vertical position of an item as number of places from the top. The second and fifth columns show how many times an item would occur in 100 selections for the small menu, while the third, sixth and eighth columns show how many times an item would occur in 100 selections for the large menu. A number of different selection frequency distributions are reported in the literature.

Table 15: Selection frequency of small and large menu items and their distribution

Item	Distribution		Item	Distribution		Item	Distribution
	Small	Large		Small	Large		Large
1	0	0	11	2	4	21	2
2	0	0	12	4	6	22	4
3	4	0	13	10	0	23	0
4	8	0	14	12	0	24	8
5	0	6	15	2	0	25	0
6	4	8	16	20	8	26	10
7	0	4	17	8	0	27	12
8	10	6	18	-	4	28	6
9	4	0	19	-	10	29	2
10	12	0	20	-	0	-	-

However, we are interested in the distributions of difficult items where the high-frequency items can be found near the bottom of the list. The distributions for small and large menus were adapted from the literature with some modification [5].

4.6.7 Independent and Dependent Variables

Independent variables are those which were controlled during the experiment, to ensure its consistency. They are:

7. Tasks: All subjects had exactly the same number of tasks (50 selections in each session) and the same selection frequency of tasks. This was ensured by following criteria developed to ensure the consistency of tasks.

8. Interactive metaphors: All subjects were required to assess the same set of interaction metaphors.

9. Learning effect: To ensure that the learning effect was controlled, only one menu was displayed at a time, in view of the fact that the position of menus might affect time (since subjects were asked to click 'Start' and then move the cursor towards the position of the menu). In addition, placing menus in different positions in the graphical user interface might have influenced the subjects' judgment of their relative importance. Finally, each subject was assigned randomly to a group and each group was randomly assigned to a different order of conditions.

10. Task criterion time: Each task had a criterion time within which it was to be completed. A task would be regarded as unsuccessfully completed if not completed within its criterion time.

11. Amount of training: Training was recorded to ensure that all subjects had the same amount. For example, all subjects were shown how to start the experiment and select the required items. In addition, we allowed subjects to ask questions before starting the experiment. Apart from this, the same amount of information was given to all subjects in all groups.

The dependent variables were grouped into matrices.

Efficiency

1. Task accomplishment time: The time taken to complete the task. Time was counted automatically during the tasks.
2. Number of errors: The number of errors occurring during each task. Errors were also counted automatically during the tasks.

Effectiveness

1. Percentage of tasks successfully completed by all subjects: This was calculated for the number of tasks completed within their criterion times.
2. Number of subjects who successfully completed all tasks: This again required all tasks to be completed within their criterion times.

Satisfaction

1. Overall satisfaction: Subject satisfaction was measured for each interaction metaphor utilised during the experiments.

Customisation behaviour

1. Customisation time: The amount of time spent on customisation. This was counted automatically during the tasks.
2. Number of clicks required to move items up or down: Clicks were counted automatically during the tasks.
3. Number of times that subjects chose recency and frequency options in the mixed-initiative condition. Choices were also counted automatically during the tasks.

4.6.8 Procedure

First, subjects were randomly assigned to different orders of conditions depending on the order of arrival, then a questionnaire was used to obtain information on user demographics, education and computer experience. Before starting each menu condition, subjects were given a recorded tutorial. In the experiment, the subjects performed the five conditions in a predetermined order given by the experimenter. First, they were asked to choose the menu condition according to the order given by the experimenter. The first task session began when the subjects clicked the 'Start' button. Next, a target item was displayed on the screen and subjects were asked to select the same item from the pull-down menu as quickly and accurately as possible. If the wrong item was clicked a cross symbol appeared on the screen. The second target item appeared once the target item had been selected. When a subject selected the correct item, the menu was disabled for 1 second before the next item. Time between the presentation of the target item and the correct selection was recorded, as well as the number of errors (incorrect selections). In the adaptable, highlighted and minimised menus, subjects were told that they could change the positions of the items if they wanted to do so after the first session. In addition, the time required by each subject to customise these menus was recorded. In session 2, item positions remained as they were at the end of session 1, unless subjects customised the positions of menu items. The primary reason for this was to measure the effects of the changes made in session 1, since subjects performed differently. In other words, if subjects had begun session 2 from the same point that they had begun session 1, the result would not have been expected to change. On the other hand, menu design remained as it was, to unify menu conditions across all sessions. For example, in highlighted and mixed-initiative menus the highlighted items would fade away. Finally, a feedback questionnaire was used to rank the menu conditions, to assess subjects' satisfaction and to record any additional comments.

4.6.9 Training

Each subject attended a five-minute recorded training session about their environment before doing the requested tasks. Additional explanation was sometimes provided when needed.

4.6.10 Data Collection

Quantitative and qualitative data were collected by recording experiments, questionnaires, interviews and observation. Experiments were not recorded, since the time taken to perform the tasks and the number of errors were automatically calculated by the application. In addition, it calculated precisely the time take to customise the menus and the frequency of clicks on the 'recently' and 'frequently' options in the mixed-initiative menu. Questionnaires and interviews also provided data on subjects' opinions and levels of satisfaction, while observation and notes taken during the experiments helped to improve understanding of each condition and to collect the required data. These measures are described below and grouped according to category.

4.6.11 Measurement

In order to fulfil the objectives mentioned in Section 4.3, three usability parameters had to be measured first (see Table 16). Efficiency can be calculated by measuring the amount of effort required to accomplish certain goals or tasks [204, 205]. Thus, efficiency was measured by the time subjects took to complete tasks and by the number of errors made during the accomplishment of each task. Effectiveness can be measured in terms of whether certain goals or tasks are achieved successfully [204, 205]. Hence, effectiveness was measured by calculating the percentage of subjects who completed tasks along with the percentage of tasks completed by all subjects. To compare the effectiveness of the five conditions a critical time for task completion was derived for each menu size (small and large). A task would then be regarded as successfully completed if subjects completed it within the critical completion time. Satisfaction was measured qualitatively by attitude rating scales, asking subjects to rate their satisfaction [204, 205].

Table 16: Metrics and dependant variables

Metric	Dependent variables
Efficiency	1.Time taken to complete the tasks 2.Number of mouse clicks 3.Number of errors
Effectiveness	1.Percentage of tasks successfully completed 2.Number of subjects who successfully completed all tasks
Satisfaction	1.Overall satisfaction
Customisation behaviour	1.Time taken to customise 2.Number of mouse clicks

4.7 Results

The experimental results comprised both quantitative and qualitative measures, along with self-reported and observed data. In addition, interviews were conducted with subjects when needed. In the case of large menus, it was noticeable that subjects who participated in the evaluation of the adaptable and highlighted conditions hesitated to customise their menus. In addition, it was apparent that subjects spent less time in customisation under the adaptable than the highlighted condition. This may be related to the fact that the boldface technique helped subjects to customise more easily.

4.7.1 Efficiency

4.7.1.1 Small Menu Selection Time

A one-way repeated measures ANOVA was performed and showed that there were significant differences in efficiency among the five menu conditions: overall at 0.05 (F (4, 236) = 30.615, $p < 0.001$) and in session 1 (F (4, 120) = 5.45, $p < 0.001$). By contrast, there were no significant differences in session 2 (F (0.016, 0.48) = 39.7, $p > 0.05$). On the other hand, the multivariate test revealed significant differences in session 2 (V = 0.96, F (4, 9) = 1.27, $p < 0.001$). When the differences between the menu types were analysed by using the t-test at 0.05 in both sessions, it was found that as expected in hypothesis 1 stated in section 4.5 in session 1 subjects were significantly faster with the adaptable menus than the adaptive minimised menu ($t29$ = 3.77, p = 0.003, r = 0.57), the mixed-initiative ($t29$ = 2.61, p = 0.01, r = 0.44) and split menus ($t29$ = 3.3, p = 0.003, r = 0.52), but there was no significant difference at 0.05 for the highlighted menu ($t29$ = 1.83, p = 0.77, r = 0.32).

Figure 24 shows that the highlighted menu was found to be the second most efficient condition in session 1.

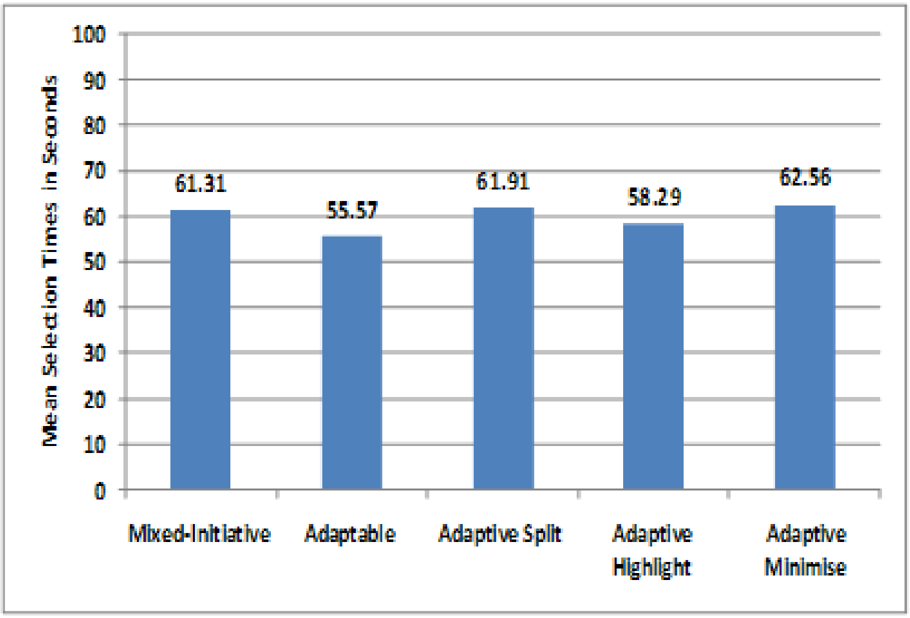


Figure 24: Mean selection times of session 1 for small menus.

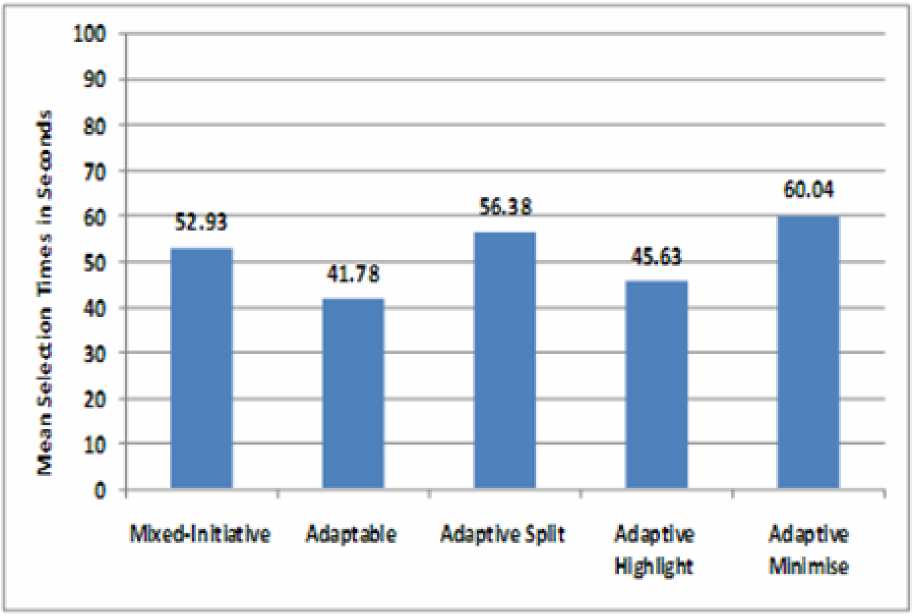


Figure 25: Mean selection times of session 2 for small menus.

The t-test revealed that in session 1 subjects were significantly faster with the highlighted menu than the minimised menu ($t29 = 2.8$, $p < 0.01$, $r = 0.46$) and the split menu ($t29 = 2.17$, $p < 0.05$, $r = 0.38$). In session 2, surprisingly, the adaptable menu was significantly faster than all other menus: adaptive minimised ($t(29) = 8.73$, $p = 0.00$, $r = 0.85$), adaptive highlighted ($t(29) = 2.51$, $p = 0.018$, $r = 0.42$), mixed-initiative ($t(29) = 8.03$, $p = 0.00$, $r = 0.83$) and adaptive split ($t(29) = 13.2$, $p = 0.00$, $r = 0.93$) (see Figure 25).

The highlighted menu was again found to be the second most efficient condition in session 2. The t-test revealed that subjects were significantly faster with the highlighted menu than the minimised menu ($t29 = 6.52$, $p < 0.01$, $r = 0.77$), mixed-initiative menu ($t(29) = 4.14$, $p = 0.00$, $r = 0.61$) and the adaptive split menu ($t(29) = 7.86$, $p = 0.00$, $r = 0.83$).

Overall, the adaptable and highlighted menus were more efficient than the others (see Figure 26). Table 17 illustrates the results of the comparison of adaptivity and adaptability levels. That by comparing between the three adaptive menus (Highlighted, Split, and Minimised) in session 1 and the comparison between the adaptable menus (Adaptable, Highlighted, and Minimised) in session 2. For the adaptivity levels the highlighted techniques was faster than the split and minimised techniques. On the other hand, for the adaptable levels, the adaptable menu was faster than the highlighted and minimised techniques.

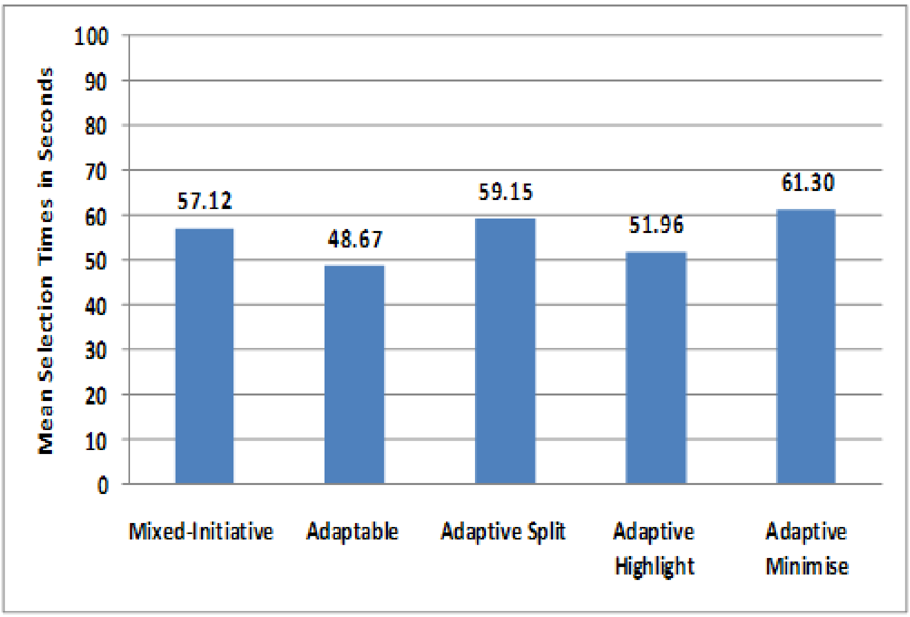


Figure 26: Mean selection times for overall for small menus.

Table 17: Comparison of adaptive menus (session 1) and adaptable menus (session 2)

	Menu efficiency		
	Most efficient	Second most efficient	Third most efficient
Adaptive menus Session 1	Highlighted	Split	Minimised
Adaptable menus Session 2	Adaptable	Highlighted	Minimised

Table 18 shows a summary of the results obtained from t-tests, with statistically significant results displayed in **bold**. MI = Mixed Initiative, AD = Adaptable, AM = Minimised, AH = Highlighted, AS = Split.

Table 18: T-test results for small menus

Conditions	Session 1	Session 2
AD vs. AM	**t29 = 3.77, p = 0.003, r = 0.57**	**t(29) = 8.73, p = 0.00, r = 0.85**
AD vs. AH	t29 = 1.83, p = 0.77, r = 0.32	**t(29) = 2.51, p = 0.018, r = 0.42**
AD vs. MI	**t29 = 2.61, p = 0.01, r = 0.44**	**t(29) = 8.03, p = 0.00, r = 0.83**
AD vs. AS	**t29 = 3.3, p = 0.003, r = 0.52**	**t(29) = 13.2, p = 0.00, r = 0.93**
MI vs. AH	t29 = 1.78, p =0.086, r = 0.31	**t(29) = 4.14, p = 0.00, r = 0.61**
MI vs. AM	t29 =0.824, p = 4.17, r = 0.15	**t(29) = 3.17, p < 0.05, r = 0.51**
AS vs. AM	t29 = 0.32, p = 0.75, r = 0.06	t(29) = 1.79, p = 0.83, r = 0.32
AS vs. AH	**t29 = 2.17, p < 0.05, r = 0.37**	**t(29) = 7.86, p = 0.00, r = 0.83**
MI vs. AS	t29 = 0.264, p =0.79, r = 0.05	**t29 = 2.74, p < 0.05, r = 0.45**
AM vs. AH	**t29 = 2.8, p < 0.01, r = 0.46**	**t29 = 6.52, p < 0.01, r = 0.77**

4.7.1.2 Large Menu Selection Time

A one-way repeated measures ANOVA showed that there were significant differences in efficiency among the five menu conditions at 0.05 overall ($F(3.24, 191.44) = 4.91, p < 0.003$) and in session 2 ($F(2.63, 76.12) = 11.30, p < 0.001$), but no significant differences were found in session 1 ($F(2.57, 74.42) = 1.90, p > 0.05$). When the differences between the menu types were analysed using the t-test it was found that in session 1 as predicted in hypothesis 2 stated in section 4.5 subjects were significantly faster with the adaptive split than the adaptable menu ($t_{29} = 2.24, p < 0.05, r = 0.38$) (see Figure 27).

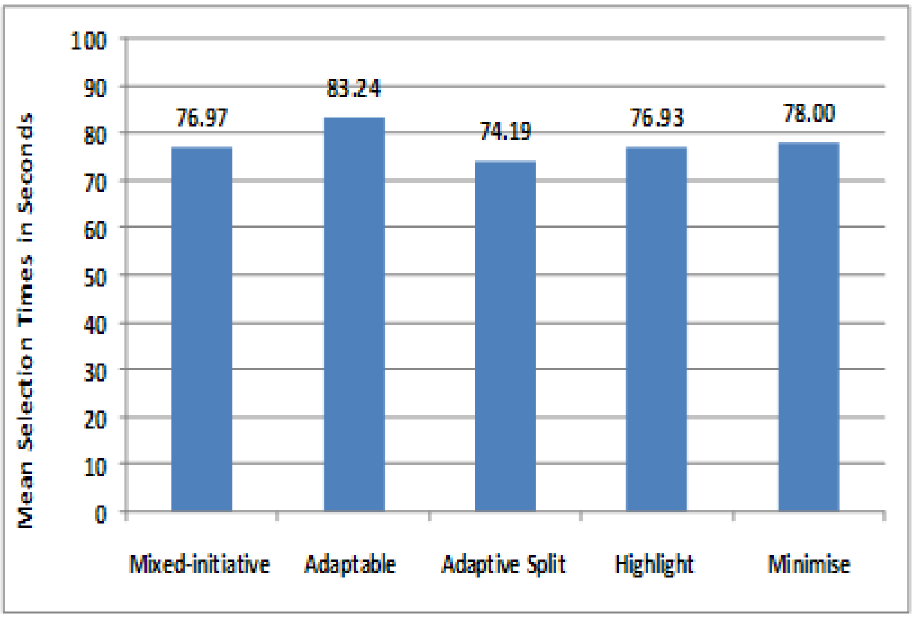


Figure 27: Mean selection times of session 1 for large menus.

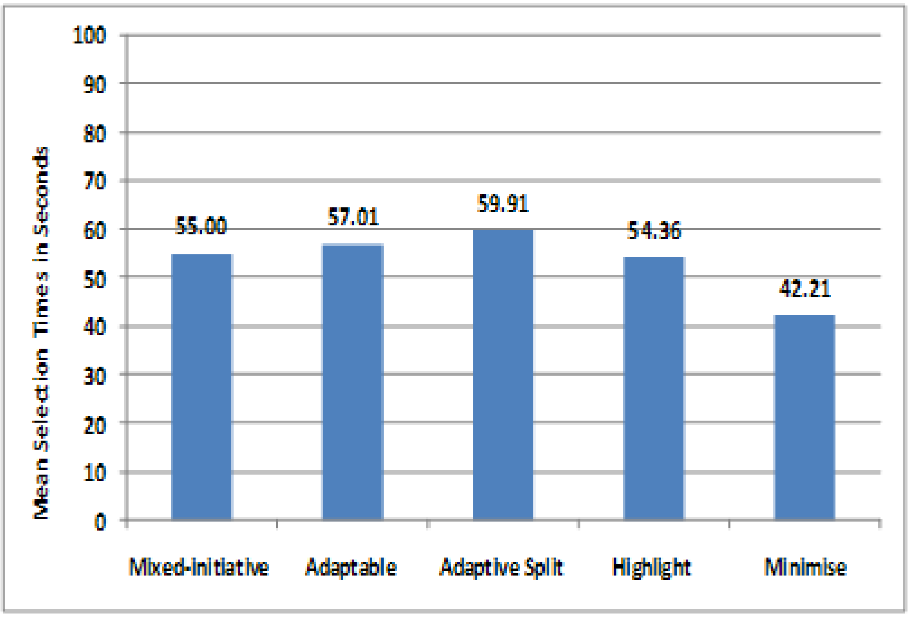


Figure 28: Mean Selection times in session 2 for large menus.

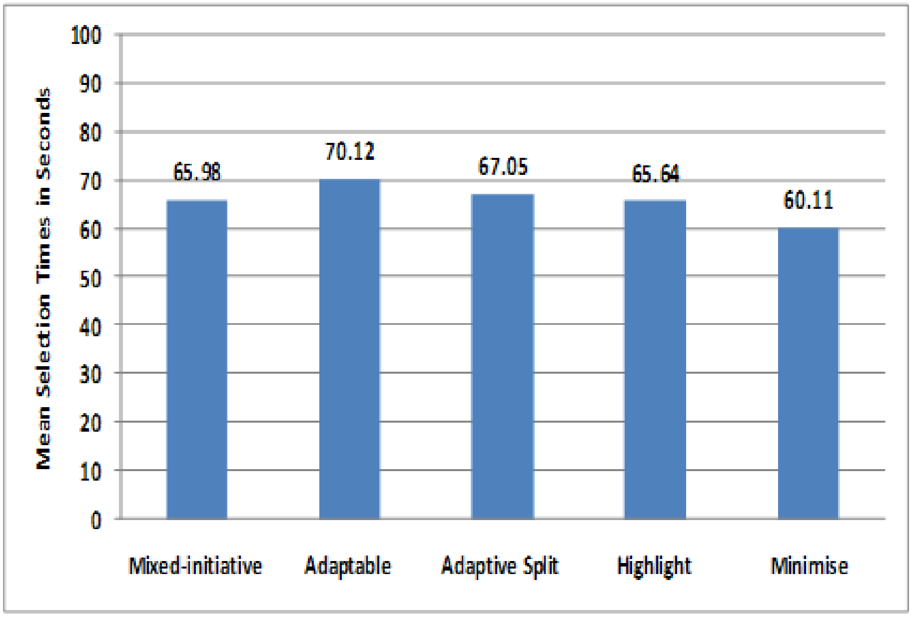


Figure 29: Mean overall selection times for large menus.

However, there were no significant differences between mixed-initiative, adaptive highlighted and minimised menus. In session 2, subjects were significantly faster with the minimised menu than the highlighted ($t29 = 5.01$, $p < 0.05$, $r = 0.68$), mixed-initiative ($t29 = 6.5$, $p < 0.05$, $r = 0.77$), split ($t29 = 9.9$, $p < 0.05$, $r = 0.87$) and adaptable menus ($t29 = 4.6$, $p < 0.05$, $r = 0.65$) (see Figure 28). Overall, the minimised menu was the most efficient (see Figure 29).

Table 19 illustrates the results of the comparison of adaptive menus in session 1 and the comparison between the adaptable menus in session 2. For the adaptivity levels the split techniques was faster than the highlighted and minimised menus. On the other hand, for the adaptable levels, the minimised technique was faster than the highlighted and adaptable techniques. Table 20 shows a summary of the results obtained from t-tests, with statistically significant ones in bold. MI = Mixed Initiative, AD = Adaptable, AM = Minimised, AH = Highlighted, AS = Split.

Table 19: Comparison of adaptive and adaptable menus

Menu	Most efficient	Second efficient	Third efficient
Adaptive menus Session 1	Split	Highlighted	Minimised
Adaptable menus Session 2	Minimised	Highlighted	Adaptable

Table 20: T-test for large menus

Conditions	Session 1	Session 2
AD vs. AM	t29 = 1.09, p = 0.287, r = 0.20	**t29 = 4.64, p < 0.01, r = 0.65**
AD vs. AH	t29 = 1.54, p = 0.133, r = 0.27	t29 = 0.77, p = 0.45, r = 0.14
AD vs. MI	t29 = 1.54, p = 0.135, r = 0.27	t29 = 0.83, p = 0.41, r = 0.02
AD vs. AS	**t29 = 2.24, p < 0.05, r = 0.38**	t29 = 0.83, p = 0.42, r = 0.02
MI vs. AH	t29 = 0.013, p = 0.99, r = 0.002	t29 = 0.19, p = 0.85, r = 0.04
MI vs. AM	t29 = 0.41, p = 0.69, r = 0.08	**t29 = 6.50, p < 0.01, r = 0.77**
AS vs. AM	t29 = 1.52, p = 0.14, r = 0.27	**t29 = 9.99, p < 0.01, r = 0.88**
AS vs. AH	t29 = 1.08, p = 0.29, r = 0.20	t29 = 1.62, p = 0.12, r = 0.29
MI vs. AS	t29 = 0.97, p = 0.34, r = 0.18	**t29 = 2.19, p = 0.04, r = 0.38**
AM vs. AH	t29 = 0.35, p = 0.73, r = 0.06	**t29 = 5.01, p < 0.01, r = 0.68**

4.7.1.3 Comparing Small and Large Menu Selection Time

A 2 x 5 (screen size x menu conditions) repeated measures ANOVA was performed. Mauchly's test indicated that the assumption of sphericity had been violated for the main effects of the menus ($X^2(2) = 16.98$, $p < .05$) and for the interaction between menu type and size ($X^2(2) = 25.47$, $p < .05$). Therefore, degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = 3.43$ for the main effect of menus and 3.24 for the main effect of the interaction between menu type and size). All effects were reported as significant at $p < .05$. There was a significant main effect of menu type on ratings of the menus: $F(3.43, 202.52) = 3.34$. Comparisons revealed that adaptable menus were faster than adaptive split menus: $F(1,59) = 5.41$, $r = 0.29$. No other significant differences were found when comparing minimised menus to the baseline (adaptable) ($F(1,59) = .73$, $r = 0.11$), highlighted menus to adaptable ($F(1,59) = .16$, $r = 0.05$) or mixed-initiative menus to adaptable ($F(1,59) = 2.01$, $r = 0.18$). In addition, there was a significant main effect of size on menu ratings ($F(1.0, 59.0) = 13.107$) and a significant interaction effect between menu type and size ($F(1, 59) = 13.11$). This indicates that size had different effects on participants' performance depending on which type of menu was used. To break down this interaction, all menu types were compared to the baseline (adaptable menu) and both sizes to the baseline (small). This revealed significant interactions when comparing small with large menus for minimised ($F(1,59) = 40.22$, $r = 0.64$), highlighted ($F(1,59) = 7.94$, $r = 0.34$), split ($F(1,59) = 22.81$, $r = 0.52$) and mixed-initiative menus ($F(1,59) = 28.47$, $r = 0.57$), against the baseline menu (adaptable).

Figure 30 shows the mean selection time for small and large menus in session 1. Subjects who utilised the small menus were significantly faster under the adaptable approach than all other menus, with an average of 55.57 seconds, whereas this approach was the slowest for large menus, with an average of 83.24 seconds. The split was significantly slower than both adaptable and highlighted approaches in small menus, but considerably faster than all others for large menus, with an average of 74.19 seconds.

Figure 31 shows the mean selection time for small and large menus in session 2. In the small menu experiment, subjects were again significantly faster when using the adaptable menu than all other conditions, whereas the best approach for the large menus in session 2 was the minimised approach, which was considerably faster than other approaches (42.21 seconds). In addition, the split menu was the slowest (59.91 s), whereas for small menus this was the second least efficient condition (56.36 s). In session 1 of the large menu experiment, the adaptable approach was the least efficient (83.24 s). By contrast, this approach was the most efficient in small menus (only 55.57 seconds). For large menus, the split was the most efficient (74.19 s), whereas for small menus this was the second least efficient condition (61.91 s), the minimised approach being the least efficient.

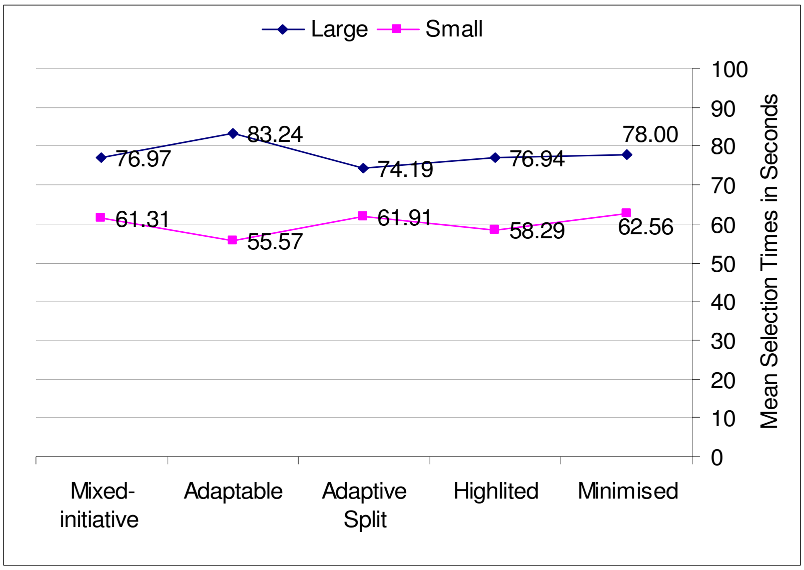


Figure 30: Mean selection times for small and large menus, session 1

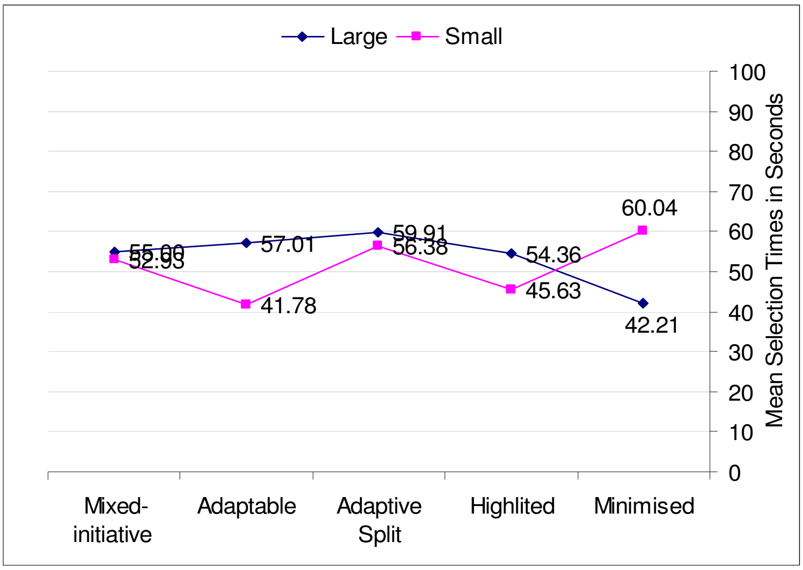


Figure 31: Mean selection times for small and large menus, session 2

On the other hand, in session 2, the adaptable condition was the most efficient approach to small menus, but the second least efficient condition in large menus (57.01 s), while the split menu was the least efficient (59.91 seconds). There was no significant difference in efficiency between the split and mixed-initiative conditions in either small or large menus. Furthermore, in session 2 of the large menu experiment, the minimised approach was surprisingly the most efficient condition, even more than its counterpart in the small menu experiment.

Figure 32 depicts graphically the overall results for both menus and shows that some approaches behaved in the opposite way to others. By way of illustration, the adaptable approach was the most efficient in small menus, but the slowest in large menus, while the opposite was true of the minimised approach, which was the most efficient in large menus and the slowest in small menus. Interestingly, there was no significant difference in this approach for the mean overall selection times between large and small menus.

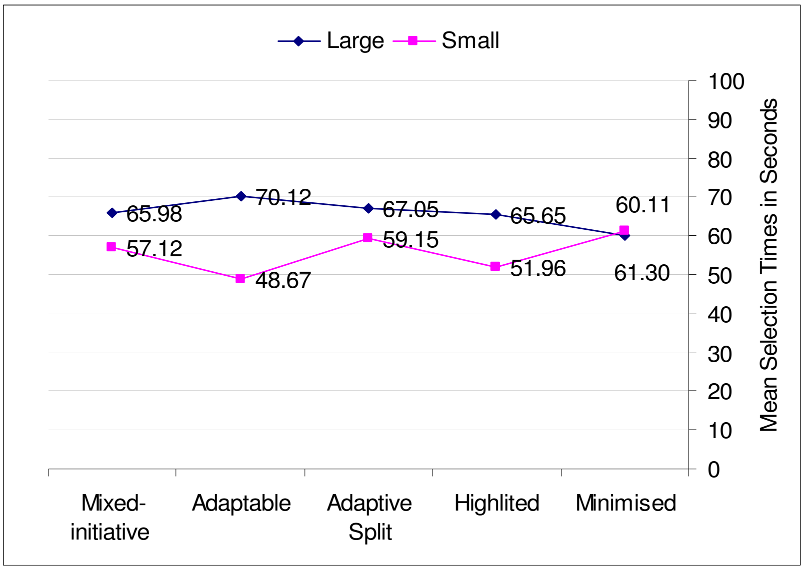


Figure 32: Mean overall selection times

4.7.2 Error Rate

Table 21 shows the total number of errors for all subjects for each menu condition. An error was recorded when a subject clicked an item that was different from the target. Each cell contains the number of errors the subjects made in 3000 selections (50 selections x 2 sessions x 30 subjects) in each condition. All the hypotheses related to frequency of error-occurrence mentioned on section 4.5 are rejected since the errors between approaches are closely similar. However, it can be seen that customisation helped to reduce the number of errors in the adaptable small menu and minimised large menu, whilst it increased between sessions 1 and 2 in the adaptable large menu. This increase is difficult to explain, but may be related to the fact that the additional content of the large menu caused confusion. In small menus, 50% of the errors were eliminated in session 2 in the adaptable approach, while improvement was slightly less marked for the highlighted menu. The mixed-initiative approach differed from all others in that the number of errors remained largely constant for both sessions and both menu sizes.

Table 21: Frequency of user errors.

Menu	Small		Large		Sum
Session	1	2	1	2	
Mixed-initiative	14	15	15	16	60
Split	11	17	13	11	52
Highlight	19	12	11	10	52
Adaptable	16	8	13	21	58
Minimised	16	16	15	7	54
Total	76	68	67	65	276

4.7.3 Effectiveness

The effectiveness of the experimental environments was measured by calculating the percentage of tasks completed successfully by all subjects and of subjects who successfully completed all tasks. As noted above, successful completion had to be within a critical time. In small menus, only one subject successfully completed all tasks within the criterion time under the adaptive split condition, whereas 4 subjects completed all tasks using mixed-initiative, highlighted and minimised menus and 5 did so using the adaptable menu. Ten tasks were successfully completed by all users within the criterion time under the mixed-initiative and minimised conditions, while the number of tasks successfully completed by all users using split, highlighted and adaptable menus were 5, 19 and 22 respectively.

In order to verify Hypotheses related to effectiveness stated in Section 4.5, the statistical t-test was used. Figure 33 shows that as expected in hypothesis 13 stated in section 4.5, for small menus, the adaptable approach was more effective than the adaptive one in terms of number of users who successfully completed all tasks. The t-test revealed that the adaptable menu was more effective than the split and the difference was significant at 0.05 ($t29 = 1.98$, $p = 0.05$, $r = 0.35$). The minimised menu was also more effective than the split and the difference was very significant at 0.05 ($t29 = 2.97$, $p = 0.006$, $r = 0.48$). Finally, the mixed-initiative approach was found to be less effective than both the adaptive and adaptable approaches in terms of number of users who successfully completed all tasks (see Table 22).

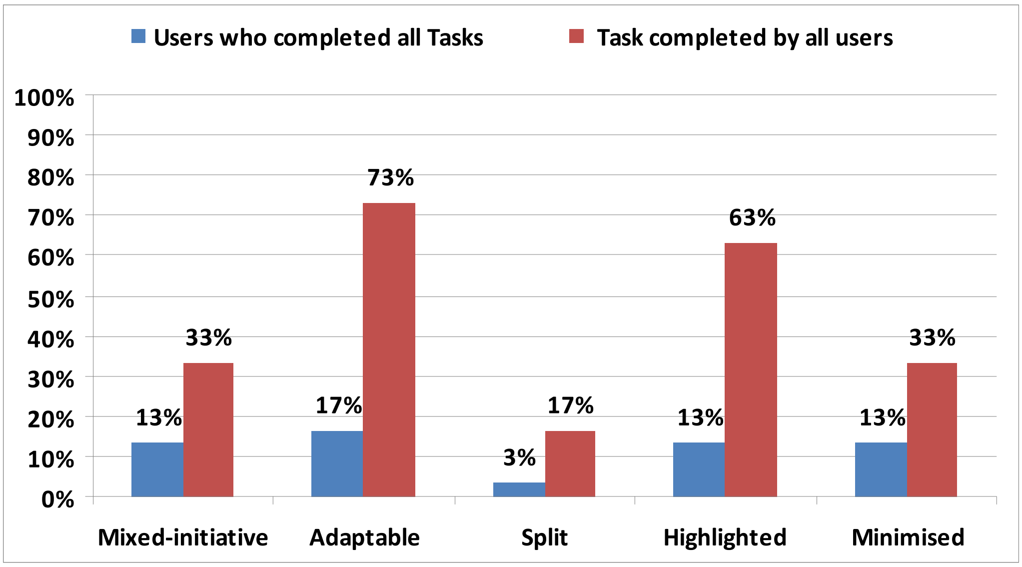


Figure 33: Effectiveness of small menus

Table 22: T-test for small menu effectiveness

Conditions	Subjects who completed all tasks	Tasks completed by all subjects
AD vs. AM	$t29 = 1.0$, $p = 0.326$, $r = 0.18$	**$t29 = 3.07$, $p = 0.005$, $r = 0.50$**
AD vs. AH	$t29 = 1.0$, $p = 0.326$, $r = 0.18$	$t29 = 1.28$, $p = 0.21$, $r =$ **0.23**
AD vs. MI	$t29 = 0.81$, $p = 0.423$, $r = 0.15$	**$t29 = 3.79$, $p = 0.001$, $r = 0.58$**
AD vs. AS	**$t29 = 1.98$, $p = 0.05$, $r = 0.35$**	**$t29 = 4.87$, $p < 0.001$, $r = 0.67$**
MI vs. AH	$t29 = 1.0$, $p = 1.0$, $r = 0.18$	**$t29 = 2.76$, $p = 0.01$, $r = 0.46$**
MI vs. AM	$t29 = 1.68$, $p = 0.1$, $r = 0.30$	$t29 = 1.0$, $p = 1.0$, $r =$ **0.18**
AS vs. AM	**$t29 = 2.97$, $p = 0.006$, $r = 0.48$**	$t29 = 1.72$, $p = 0.096$, $r =$ **0.30**
AS vs. AH	$t29 = 1.14$, $p = 0.264$, $r = 0.21$	**$t29 = 4.065$, $p < 0.01$, $r = 0.60$**
MI vs. AS	$t29 = 1.14$, $p = 0.264$, $r = 0.21$	$t29 = 1.72$, $p = 0.09$, $r =$ **0.30**
AM vs. AH	$t29 = 1.68$, $p = 0.103$, $r = 0.30$	**$t29 = 2.34$, $p = 0.03$, $r = 0.40$**

Figure 33 shows that as expected in Hypotheses 5 stated in Section 4.5, in small menus, the adaptable approach was more effective than the adaptive one in terms of number of tasks completed successfully. The t-test revealed that the adaptable menu was more effective than the split and mixed-initiative menus, the differences being very significant at 0.05 ($t29 = 4.87$, $p < 0.001$, $r = 0.67$ and $t29 = 3.79$, $p = 0.001$, $r = 0.58$ respectively). In addition, the adaptable menu was more effective than the minimised menu and the difference was significant at 0.05 ($t29 = 3.07$, $p = 0.005$, $r = 0.5$). Finally, the highlighted menu was more effective than the mixed-initiative, split and minimised menus and the difference was very significant at 0.05 ($t29 = 2.76$, $p = 0.01$, $r = 0.46$; $t29 = 4.065$, $p < 0.01$, $r = 0.6$ and $t29 = 2.34$, $p = 0.03$, $r = 0.4$ respectively).

In large menus, just 2 subjects successfully completed all tasks within the criterion time under the minimised condition, whereas 6 subjects completed all tasks using mixed-initiative and highlighted menus, 4 subjects did so using the adaptable menu and 9 did so using the split menu. On the other measure, 24 tasks were successfully completed by all users of the mixed-initiative and adaptable menus, whereas the number of task successfully completed by all users of the split, highlighted and minimised menus was 27, 29 and 26 respectively.

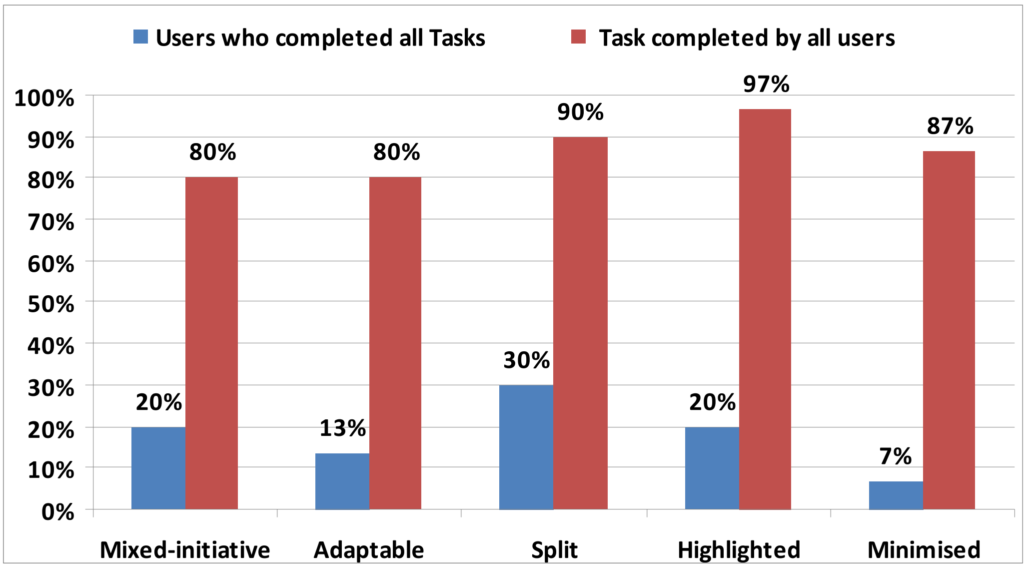


Figure 34: Effectiveness of large menus

Figure 34 shows that for large menus, the adaptable approach was more effective than the adaptive one in terms of number of users who successfully completed all tasks. In addition, the t-test revealed that as expected in hypothesis 14 stated in section 4.5 the adaptive menu was more effective than the adaptable one and the difference was significant at 0.05 (t29 = 1.98, p = 0.05, r = 0.35). The split menu was also more effective than the minimised menu and the difference here was very significant at 0.05 (t29 = 2.97, p = 0.006, r = 0.48). The mixed-initiative approach was found to be less effective than both the adaptive and adaptable approaches in terms of number of users who successfully completed all tasks (see Table 23). The statistical t-test was also used in order to verify Hypotheses 5, 6 and 7 stated in Section 4.5, revealing no differences at 0.05 for any large menus.

Table 23: T-test for effectiveness of large menus

Conditions	Subjects who completed all tasks	Tasks completed by all subjects
AD vs. AM	t29 = 1.0, p = 0.33, r = 0.18	t29 = 0.63, p = 0.54, r = 0.12
AD vs. AH	t29 = 1.0, p = 0.33, r = 0.18	t29 = 1.54, p = 0.14, r = 0.28
AD vs. MI	t29 = 8.1, p = 0.42, r = 0.83	t29 = 1.0, p = 1.0, r = 0.18
AD vs. AS	**t29 = 1.98, p = 0.05, r = 0.35**	t29 = 0.53, p = 0.60, r = 0.10
MI vs. AH	t29 = 1.0, p = 1.0, r = 0.18	t29 = 1.72, p = 0.96, r = 0.30
MI vs. AM	t29 = 1.68, p = 0.103, r = 0.30	t29 = 0.57, p = 0.57, r = 0.11
AS vs. AM	**t29 = 2.97, p = 0.006, r = 0.48**	t29 = 1.0, p = 1.0, r = 0.18
AS vs. AH	t29 = 1.14, p = 0.26, r = 0.21	t29 = 0.68, p = 0.50, r = 0.13
MI vs. AS	t29 = 1.14, p = 0.26, r = 0.21	t29 = 0.46, p = 0.65, r = 0.09
AM vs. AH	t29 = 1.68, p = 0.103, r = 0.30	t29 = 1.0, p = 1.0, r = 0.18

Overall, for small menus, both the number of subjects who completed all tasks and the number of tasks completed by all subjects was higher under the adaptable condition than any other condition, while scores on both measures were lowest for the split condition. On the other hand, for large menus alone, the number of subjects who completed all tasks was highest for the split condition and the number of tasks completed was highest for the highlighted condition, while fewer subjects completed all tasks using the minimised menu than other types and the number of tasks completed by all subjects was lowest under the adaptable and mixed-initiative conditions. This leads us to conclude that in small menus the adaptable approach was more effective than other conditions, whereas in large menus the adaptive approach, such as highlighted and split menus, was more effective than other conditions.

4.7.4 Customisation Behaviour

Subjects were not allowed to customise during the tasks; they had one opportunity to do so before starting session 2. It was found that they spent significantly less time customising the small menus than the large ones: 1 hour and eighty two minutes and two hours twenty two minutes respectively (see Figure 35).

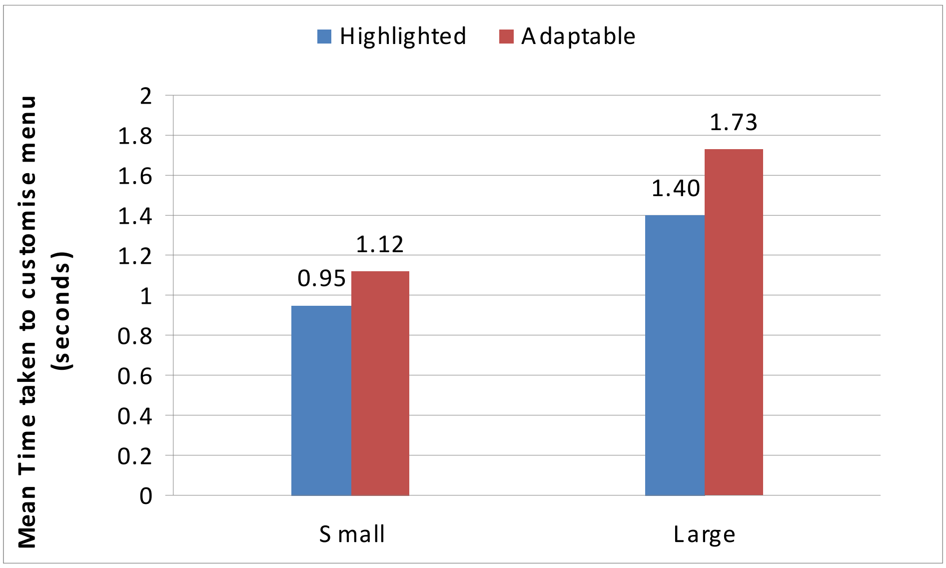


Figure 35: Customisation time

4.7.4.1 Adaptation by Users

Subjects could easily move items up and down by clicking on the required item and then on an up or down arrow placed above the menu. They were told how to customise and provided with help when needed, since we were interested in the results of customisation, not the way in which it was done. However, it was observed that subjects utilised different criteria for ordering the menu items. The most common approaches were frequency-based and alphabetical ordering. This did not prevent some subjects from using their own criteria. For example, one subject moved the items near the top to the top of the list and items near the bottom to the bottom of the list. In the small menu experiment, the results for the adaptable menu show that it was more efficient than both the highlighted and minimised menus in session 2. In addition, the highlighted and minimised menus were approximately the same in session 2. It is difficult to explain this result, but it may be related to the fact that when asked in the interview, subjects said that they felt in more control when using the adaptable menu than either the highlighted or minimised ones. For large menus, the minimised type was found to be more efficient than both highlighted and adaptable menus.

The results show that subjects behaved differently towards highlighted and adaptable menus according to their size; for example, they customised large menus less than small ones. In addition, subjects who customised adaptable menus spent more time on large than small ones, while those who customised highlighted menus spent less time on large than small ones. These results may be explained by the fact that highlighting some of the items in a small menu makes it appear visually more complex than highlighting the same number in a larger one. It was also found that under the mixed-initiative condition, subjects utilised the frequency and recency techniques more in large than small menus, with respective totals of 120 and 94 selections made by subjects.

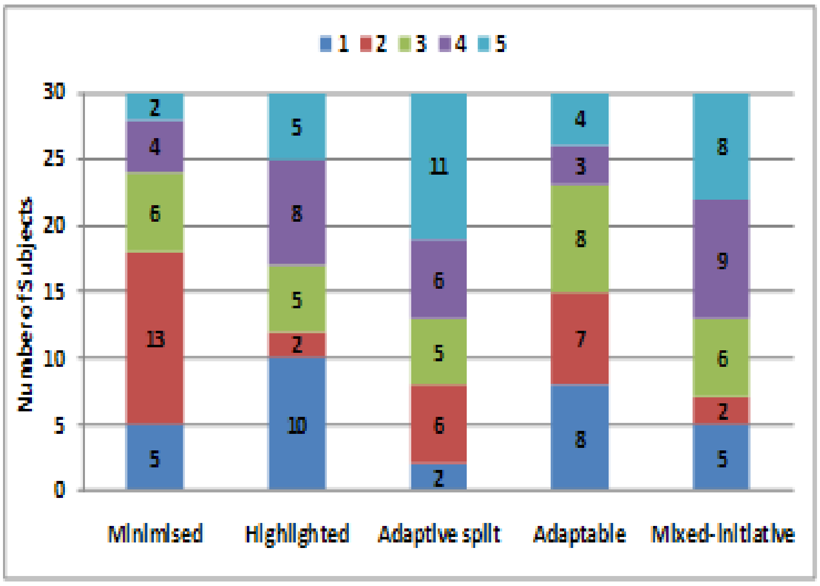
4.7.4.2 Adaptation by the System

A second objective, as mentioned early in this chapter (Section 4.3), was to investigate the effect of different levels of adaptation. Therefore, we conducted a comparison of three adaptive menus in session 1 presented with different types of adaptation: (1) changes occurring without moving items (that is, highlighted menu), (2) changes made by moving recently and frequently clicked items to the top of the list and leaving the others unchanged (that is, split menu) and (3) changes made by moving only frequently clicked items to the top of the list and leaving the others unchanged (that is, minimised menu). On one hand, the results for the adaptive menus in small conditions show that highlighting the frequently clicked items was more efficient than changing both the recently and frequently clicked items or solely the frequently clicked items. This result may be explained by the fact that subjects preferred fewer changes to occur. On the other hand, the results for adaptive menus showed that changing both the recently and frequently clicked items was more efficient than just changing or highlighting the frequently clicked items.

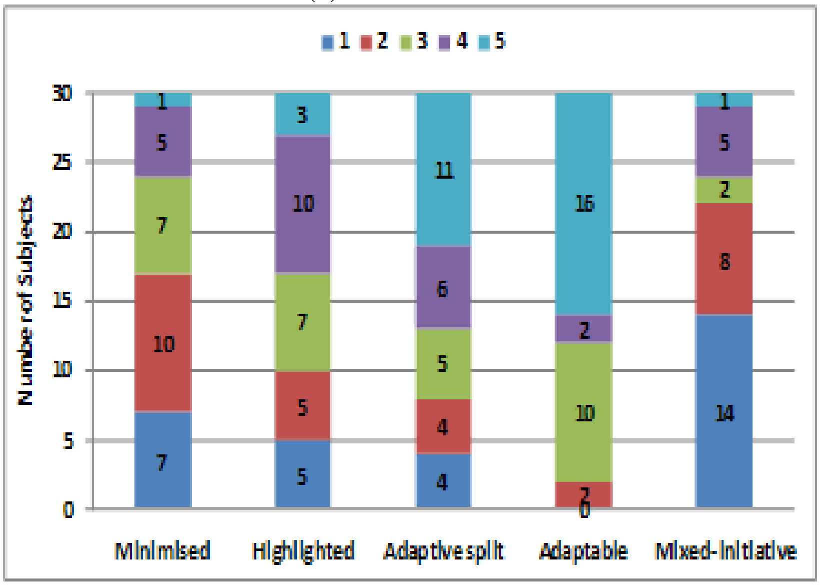
4.7.5 User Satisfaction

At the end of the experiment subjects were asked to give ratings on a 1 to 5 scale of user preferences. As Figure 36 (a) shows, for small menus exactly one-third of subjects selected the highlighted menu as the preferred type, followed by the adaptable with 8 subjects, while exactly the same number of subjects (5) chose both the minimised and mixed-initiative menus as the best menu. By contrast, only two subjects selected the adaptable menu as the best. On the other hand, the split menu was ranked last by the greatest number of subjects (just over one-third), followed by the mixed-initiative, highlighted and adaptable menus with 8, 5, and 4 subjects respectively. Only two subjects rated the minimised menu lowest. For large menus, Figure 36 (b) shows that just under the half of subjects preferred the mixed-initiative menu, followed by the minimised, highlighted and split menus with 7, 5 and 4 subjects respectively, while no subject selected the adaptable menu as the best. Indeed, this menu was categorised by 16 subjects as the most undesirable, followed by the split with 11 subjects.

Overall, for small menus, the minimised type was the most strongly preferred, as more than half of subjects ranked it either first or second. This was followed by the adaptable and highlighted menus with exactly half and more than one-third of subjects ranking them first or second respectively. The least strongly preferred types (by 17 subjects each) were the split and mixed-initiative menus. However, 11 of the 17 subjects ranked the split menu last, while 8 did so for the mixed-initiative type, suggesting that the former was the less preferred of the two. On the other hand, for large menus, the mixed-initiative type was ranked first by thirteen subjects, with more than two-thirds ranking it either first or second, followed by the minimised menu, ranked first or second by more than half of subjects. The least desirable was the adaptable type, followed by the split menu.



(a) Small menus



(b) Large menus

Figure 36: Satisfaction for small and large menus.

4.8 Discussion

The aim of this chapter is to compare the three personalisation approaches: adaptive, adaptable and mixed-initiative. For small menus, the results show that the adaptable menu was the fastest in both sessions (see Figures 24 and 25), from which we may conclude that for limited content, adaptable menus are more efficient than adaptive ones. A possible explanation for this is that the size of the menu helps subjects to remember the position of items. This can be confirmed by the observation that some subjects preferred not to customise the adaptable menus. This finding is in agreement with those of Findlater and McGrenere [7] and of Park et al. [1], who report that adaptable menus were more efficient than adaptive ones. Although these results differ from some published studies (e.g. Sears and Shneiderman [113]), they are consistent with the finding that a split menu was faster than a highlighted one. Moreover, in mixed-initiative menus, there were two reasons for subjects' uncertainty. First, the mixed-initiative menu repeatedly updates the items in the recently-used list. The second reason is that subjects need to choose to display either the recently or frequently-used items. By contrast, in the minimised menu the recency technique is neglected and only the frequency is taken into account. These drawbacks seem to limit the effectiveness of both menu types. In large menus, the results show that the split menu was the fastest in session 1, although it was surprisingly the slowest in session 2. This leads us to conclude that when content is large using adaptation as a regular technique might strain subjects. This finding is in agreement with that of Sears and Shneiderman [113], who report that the split menu was faster than the adaptable one. Although these results also differ from some published studies (e.g. Findlater and McGrenere [7]), they are consistent with the finding that adaptable menus were more efficient than adaptive ones. Importantly, some menus utilised different approaches from one session to the other. For example, the adaptable menu employed the traditional approach in session 1, as subjects did not adapt it until the second session. Similarly, the highlighted and minimised menus utilised adaptive techniques in session 1 and adaptable ones in session 2. The difference here is that in highlighted menus, item positions remain the same, while in minimised ones they change. As for the mixed-initiative and split menus, these maintained the same approach in both sessions.

One of the main usability parameters is user satisfaction. In this experiment we attempted to assess which personalisation approach was preferred by users. In addition, this study examined users' views of the amount of personalisation (adaptive and adaptability levels). More specifically, it examined whether the size of personalised menus affected user satisfaction. The results indicate that user satisfaction is affected by the size of personalised content, since satisfaction varied according to size of content. For example, in large menus, the mixed-initiative menu was the most strongly preferred, while in small menus it was the least strongly preferred. In addition, the adaptable menu was the second most strongly preferred menu, while in small format it was the most strongly rejected menu. These results may be explained by the fact that in large menus users prefer to have less control and to have more help from the system, since large content requires more effort and user attention, whereas in small menus, they prefer to have full control because this control will not require much effort. This indicates that users prefer to have control as long as it does not require too much attention and effort.

There was a variety of response towards the design of each approach. First, in terms of design of the menu, subject generally liked the way that the system assisted them by moving items to the top and hiding unwanted ones. However, there were comments suggesting that the possibility of undoing the adaptation action is essential. In other words, there was a need to employ adaptation but with less movement. On the other hand, in terms of design of the adaptable menu, subjects generally liked the method of moving menu items up and down. However, in terms of design of the mixed-initiative menu, subject generally liked the chosen techniques and the recommendations provided by the system. This confirmed that the mixed-initiative approach was generally acceptable. Ultimately, during the experiment it was noticeable that subjects were willing to accept suggestions from the system while performing their tasks. In terms of design of the split menu, subjects generally showed that they did not understand the method of moving menu items up and down. In addition, there were comments suggesting that moving items continually was confusing. In terms of design of the highlighted menu, subjects generally liked the technique of boldfacing the most frequently selected items, rather than moving items.

4.8.1 Adaptable Menu

The traditional approach in session 1 was based on users memorising item positions; as pointed out by [1], it takes time to memorise the position of all the items and even when the position of frequently used items is known, the menu does not provide any support. The results for this approach varied according to menu size: it was the fastest condition for small menus but the slowest for large ones. This confirms that the traditional approach is efficient for small menus but less so as content increases. In session 2, subjects were able to customise the menu by reordering the items or putting frequently used ones at the top of the list. However, they still had to memorise the positions of items in order to customise the menu. Again, this approach was the fastest for small menus, while for large ones it was found to be the second slowest.

4.8.2 Highlighted Menus

The highlighted approach required less memorising of item positions, since the menus provided support by highlighting the position of frequently used items. The results show that because the frequently used items were already known in small highlighted menus, subjects took slightly less time to customise the menu: an average of 8.59 minutes, compared to 8.89 minutes for adaptable menus. The difference was much greater with large menus: an average of 5.52 minutes compared to 11.67 minutes for the adaptable condition. However, in small menus, the highlighted type had no significant advantage over the adaptable one in either session in terms of selection time, whereas in large menus this condition had an advantage over adaptable menus in both sessions. This confirms that the highlighting technique becomes more efficient as content increases.

4.8.3 Adaptive Split Menus

In large menus, the split technique was faster than other conditions in session 1, but was surprisingly the slowest in session 2. A possible explanation for this is that the size of the searching area affected subjects' behaviour, since in session 1 they had to consider the whole menu, whereas in session 2 the frequently clicked items moved to the top of the list and subjects neglected the bottom of the menu, which narrowed the search area. This can be confirmed by observation and interviewing subjects after the experiment. In addition, the results for small menus in session 1 show that the split type was very slow. This result is consistent with those of Findlater and McGrenere [5], who report that accessing menu items on a small screen was slower than on a large one. This may also be the case for searching for items among small content compared to a large one.

4.8.4 Adaptive Minimised Menus

In the minimised menu the recency technique was neglected and only the frequency was taken into account. This design caused subjects to obtain the benefit of the frequently used items only, which seems to have limited the effectiveness of this menu. Therefore, further work needs to be done to establish whether utilising both recency and frequency techniques would be more beneficial.

4.8.5 Mixed-initiative Menus

In mixed-initiative menus there were two reasons for uncertainty among subjects: first, this type of menu repeatedly updates the items in the recently-used list; secondly, subjects must choose to display either recently or frequently-used items. These drawbacks seem to limit the effectiveness of this menu type. There is therefore a definite need to show both recently and frequently used items, while avoiding repeated updates of the items. Therefore, further work needs to be done to establish whether utilising both the recency and frequency techniques would be more beneficial.

4.9 Summary

This chapter has documented the second set of experiments carried out to investigate the efficiency, effectiveness and satisfaction of users of different types of personalised menus, as measures of usability. The metrics used to measure efficiency were task accomplishment time and frequency, while effectiveness was measured by calculating the number of subjects completing all tasks in sessions 1 and 2, and the number of tasks completed successfully within task criterion times. Satisfaction was measured by using 5-point Likert scales.

For small menus, results show that overall the adaptable menu was surprisingly the best in terms of efficiency. Errors were also reduced in the adaptable menu by 50% when subjects customised their menus. Unexpectedly, subjects were slower using the split, mixed-initiative and minimised menus. For large menus, the split menu condition was found to be the best overall in terms of efficiency on initial use, while the adaptable type was slower than the highlighted and mixed-initiative menus. The minimised menu was found to be the best in terms of efficiency for the second time of use. A comparison between small and large menus shows that the adaptable type was surprisingly the most efficient overall of the small menus and the least efficient of the large ones. Conversely, the minimised type was the slowest of the small menus and the fastest of the large ones. Finally, errors were reduced in adaptable and minimised small menus by 50% and 62% respectively, whilst being increased in the large adaptable type. As mentioned in section 4.5, the design of the adaptive minimised menus caused subjects to obtain the benefit of the frequently used items only, which seems to have limited the effectiveness of this type. Therefore, further work needs to be done to establish whether utilising both recency and frequency techniques would be more beneficial. In addition, the mixed-initiative menu repeatedly updates the items in the recently-used list and subjects must choose to display either recently or frequently-used items. These drawbacks seem to limit the effectiveness of these two menu types. There is therefore a clear need to show both recently and frequently used items, while avoiding repeated updates of the items.

Finally, it cannot be fully concluded that an adaptable menus was more or less efficient, effective and satisfactory than the counterpart of mixed-initiative and adaptive menus in either menus size (small or large menus), since these two conditions have design limitations. Therefore, the three main interaction approaches (adaptive, adaptable and mixed-initiative) will be further investigated. More specifically, the next chapter will seek the design enhancement of these two menu types, along with an enhancement of the adaptable condition.

Chapter 5: Study Three: Comparison of Improved Personalisation Approaches

5.1 Introduction

From the second experiment it was noticeable that there were design limitations to the adaptive/adaptable minimised and mixed-initiative menus. As expected, in the adaptable condition, subjects could not customise their menus unless they had used them for some time, since they would not know what changes to make. These drawbacks affected the efficiency, effectiveness and satisfaction levels of the adaptive, adaptable and mixed-initiative approaches. Therefore, multimodal auditory solutions (speech, earcons and auditory icons) were utilised in order to mitigate the drawbacks of these menus, improve their performance and increase their usability.

This chapter begins with a statement of the aims and objectives of the third experiment, followed by a description of the experimental platform. The hypotheses are then stated and the experimental design and methodology are described. Finally, the chapter presents an illustrative and descriptive analysis of the data obtained and discusses the results of the three usability parameters (efficiency, effectiveness and satisfaction) along with subjects' customisation behaviour (see Figure 37).

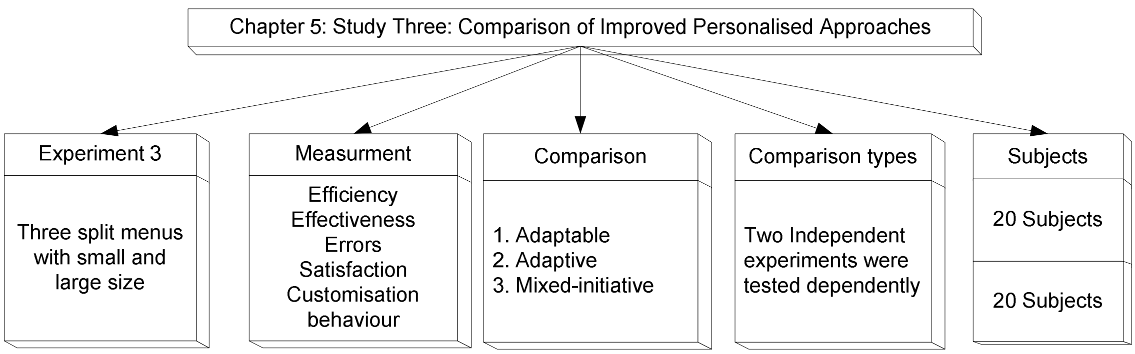


Figure 37: Summary of chapter 5

5.2 Aims

The aims of the third experiment were first to improve the adaptive, mixed-initiative and adaptable menus and then to establish which condition was most usable, since the second experiment had shown that there were design limitations to these menus that might affect their usability. It also aimed to investigate the effects of providing an 'as you go' incremental customisation strategy during the tasks. More specifically, it compared the performance of the adaptable menu when customised during the task using speech against the same menu without that customisation.

5.3 Objectives

In order to fulfil our goals, four objectives had to be attained. The first was to measure precisely the efficiency of each condition by timing the completion of tasks and counting the errors made in each condition; the second objective was to measure the effectiveness of each condition by calculating the percentage of tasks completed successfully; the third was to measure user satisfaction under each condition in small and large menus by obtaining subjects' opinions; and the fourth was to examine subjects' customisation behaviour using speech.

5.4 Experimental Approaches

The layout utilised here, shown in Figure 38, was the same as that used in the second experiment. In addition, to unify the design of the menus, only the split technique was used, since this was the most appropriate way to display both recently and frequently selected items.

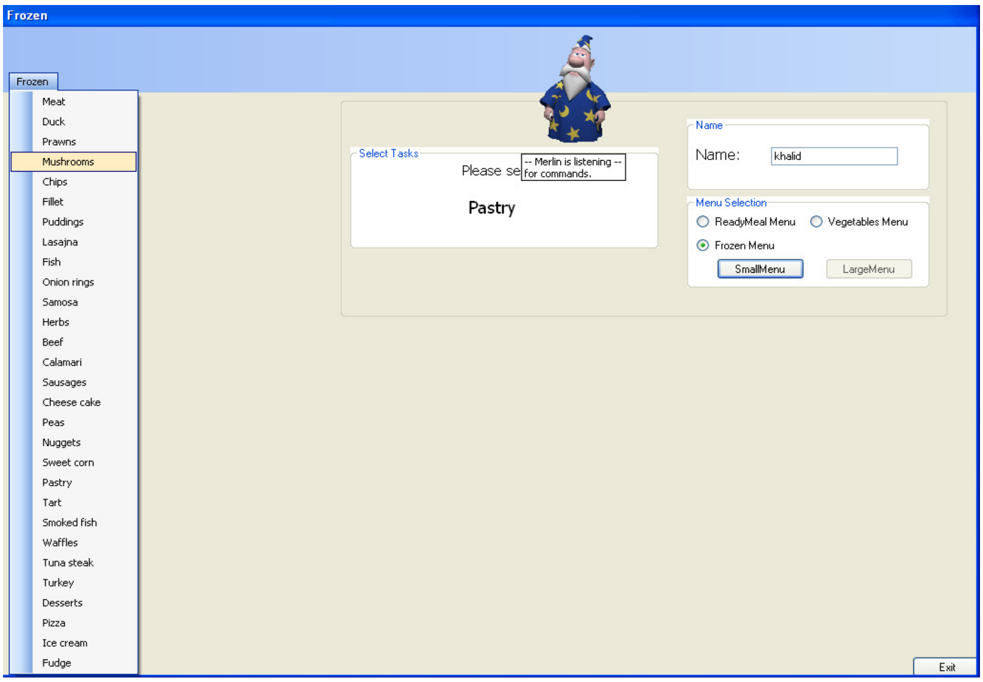


Figure 38: Screen layout utilised in third experiment

5.5 Menu Types and Designs

Three menu conditions were tested: adaptive, adaptable and mixed-initiative. Each was implemented as a split menu, in which the top section was extendable and kept the most frequently selected items separate from the bottom section, while the second section, which was also extendable, kept the most recently selected items separate from the top and bottom sections. All menus started without horizontal split lines, then once there was a recently or frequently selected item, the menu would move it and create a separate section for it by inserting a horizontal line. In addition, multimodal auditory solutions (speech, earcons and auditory icons) were utilised in order to mitigate the drawbacks of these menus, improve their performance and increase their usability.

5.5.1 Adaptive Menu

In the adaptive/adaptable minimised menu, the recency technique was neglected and only the frequency was taken into account. This design allowed subjects to obtain the benefit of the frequently used items only, which seems to have limited its effectiveness. The solution was to apply both the recency and frequency techniques, whereby the software counted how many times each item had been used, moving both recently and frequently selected items up the list and separating them from other items by horizontal lines. The top section was extendable and kept the most frequently selected items separate from the bottom section, while the second section, which was also extendable, kept the most recently selected items separate from the top and bottom sections. However, since the first item to be clicked would be considered recently selected and not yet frequently selected, the top section when first created would be for the recently selected items (Figure 39 b). Then, once there were some frequently selected items, they would move to the top section and be followed by the recently selected items (Figure 39 c).

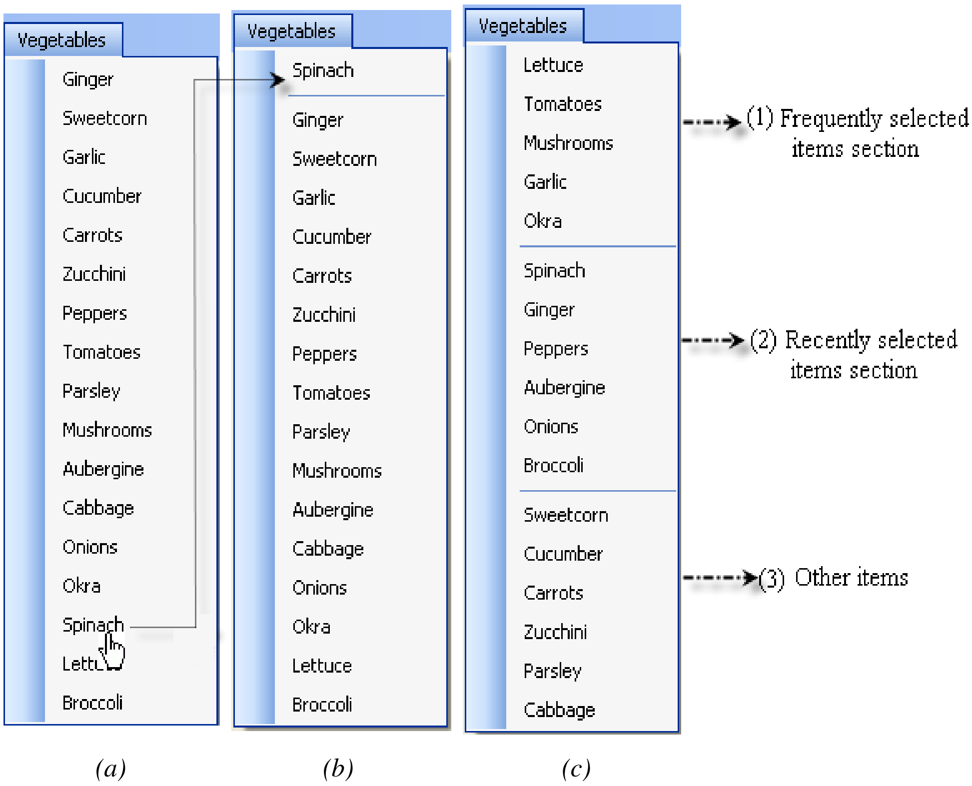


Figure 39: Adaptive menu

5.5.2 Mixed-initiative Menu

In the mixed-initiative menu there were two reasons for uncertainty among subjects: first, this type of menu repeatedly updates the items in the recently-used list; secondly, subjects must choose to display either recently or frequently-used items. In addition, it was noticeable during the second experiment that some subjects (8) did not notice the option of taking recommendations from the system. These drawbacks seem to have limited the effectiveness of this menu type. There is therefore a definite need to show both recently and frequently used items, while avoiding repeated updates of the items; the split technique is useful to solve these problems, since it allows the recently and frequently used items to be split into sections and displayed at the same time. Therefore, in the mixed-initiative menu, the technique was to display the recently or frequently used items to subjects at the appropriate time. The recently selected items were displayed at the top of the menu when this feature was selected by clicking on a button labelled ‘Recently’ (Figure 40 (2)), while the frequently selected items were displayed when the ‘Frequently’ button was selected (Figure 40 (3)). The software counted how many times each item had been used, moving both recently and frequently selected items to the top of the list and separating them from other items by a horizontal line. The recency technique moved the recently selected items to the top and once there were some frequently selected items, these would be moved underneath the top section and separated by horizontal lines from the sections above and below. Conversely, the frequency technique moved frequently selected items to the top and recently selected ones underneath, again separating the three sections with two horizontal lines. Subjects could switch from one to another at any time during the experiment; it was their responsibility to choose the appropriate technique. Furthermore, we utilised an earcon to inform users that there was a new recommendation from the system (Table 24), because in the second experiment most of the subjects (17 using large menus and 11 using small ones) had forgotten to choose the frequency technique.

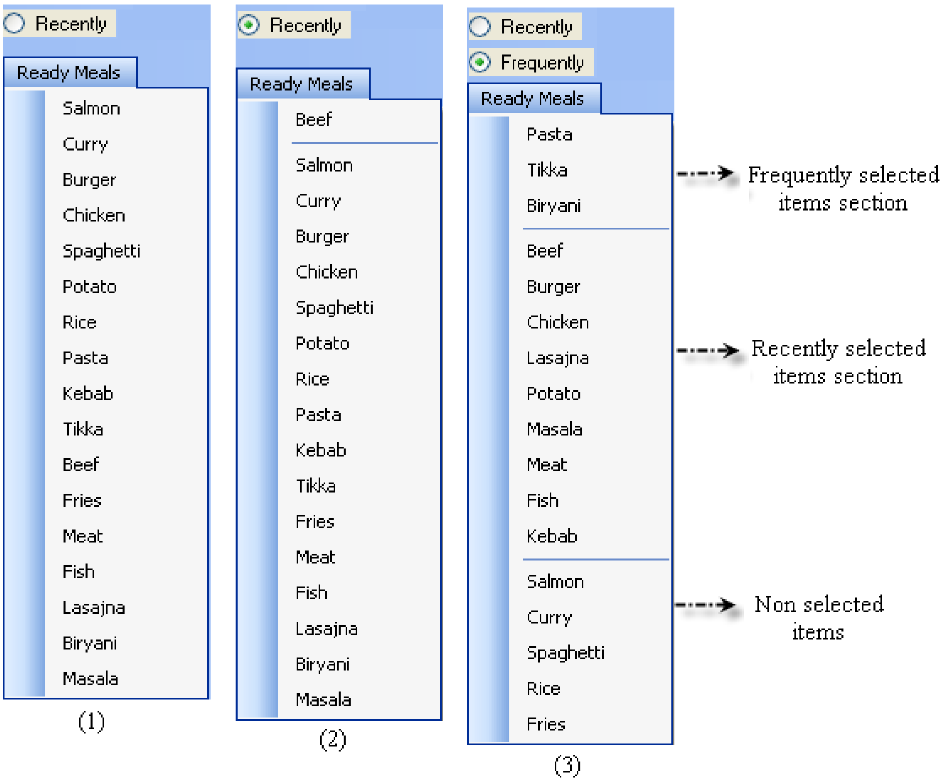


Figure 40: Mixed-initiative menu

Table 24: Earcons used to solve usability problems with menus

Condition	Rhythm	Timbre	Duration	Where used	Perceived meaning	Why used
Mixed-initiative	Piano	Rising pitch	30 ms	Recommendation	A frequency option has appeared	Subjects forgot to select this option
Adaptable	Piano	Rising pitch	50 ms	Customising menu by sound	Menu has successfully moved the item	So that subjects would know that items had moved

5.5.3 Adaptable Menu

In the adaptable condition, subjects could modify the order of items during the tasks, using speech (Figure 41). Subjects had to say the correct item name, since this study was concerned with the effects of customisation strategies, not with the means and tools utilised. The fact that we utilised an inferior speech recognition software package may have led to speech recognition problems. Furthermore, some subjects were non-native speakers and they required more training. In order to eliminate any effect of these factors on the results, it was decided that if a subject said the name of an item twice and there was no response from the system, the experimenter would move the item on behalf of the subject. This was done by using a separate keyboard, in order not to interrupt the subjects. These steps will have helped to reduce the impact of speech-recognition errors that many new tools have overcome or reduced. We also utilised earcons to inform subjects that selected items had moved, in order to help them to improve their performance. Details are given in Table 24.

Figure 41: Adaptable menu

5.6 Menu Size

Two menu sizes were used, as in the second experiment (see Chapter 4, section 4.4.2), in order to determine the effect of size on usability. The large menu was a full length one displayed on a large screen, containing 29 items of which 14 were included in the experimental tasks. The small menu was the size of many menus that are commonly used and was the minimum length that would allow the same number of items (14) to be included in the tasks as for the large one. In addition, it was approximately half the length of the larger menu.

5.7 Experimental Hypotheses

The aim of the third experiment was to measure the efficiency and user satisfaction of the adaptive, adaptable, and mixed-initiative approaches to the customisation of menus. Based on the literature review of related work, we sought to test the following hypotheses.

H8: In large menus, the adaptive approach will be more efficient than the adaptable one in terms of task accomplishment time, frequency of errors, number of tasks completed successfully, and user satisfaction.

H9: In small and large menus, the mixed-initiative approach will be more efficient than both the adaptive and adaptable conditions in terms of task accomplishment time, frequency of errors, number of tasks completed successfully, and user satisfaction.

H10: In small menus, the adaptable menu customised by sound will be more efficient than the same menu not customised by sound in terms of task accomplishment time.

H11: In large menus, the adaptable menu customised by sound will be more efficient than the same menu not customised by sound in terms of task accomplishment time.

5.8 Experimental Design

The experimental design, tasks, variables (independent and dependent variables) and measures were similar to those reported in Chapter 4, except for the procedure, training and sample size.

5.8.1 Procedure

First, subjects were randomly assigned to different orders of conditions depending on the order of arrival, then a questionnaire was used to obtain information on user demographics, education and computer experience. Before starting each menu condition, subjects were given a recorded tutorial and training session on how to customise using sound. In the experiment, the subjects performed the three conditions in a predetermined order given by the experimenter. First, they were asked to select the menu condition according to that order. The first task session began when the subjects clicked the 'Start' button. A target item was displayed on the screen and subjects were asked to select the same item from the pull-down menu as quickly and accurately as possible. If the wrong item was clicked a cross symbol appeared on the screen. The second target item appeared once the target item had been selected. When a subject selected the correct item, the menu was disabled for one second before the next item. Time between the presentation of the target item and the correct selection was recorded, as well as the number of errors (incorrect selections). In the adaptable condition, subjects were told that they could change the positions of the items during the experiment if they wanted to do so. Each subject performed one session of 50 selections for each menu; as they were able to customise the menu on session 1, there was no need to do a second session. Finally, a feedback questionnaire was used to rank the menu conditions, to assess subjects' satisfaction and to record any additional comments.

5.8.2 Training

Each subject attended a five-minute recorded training session about their environment before doing the requested tasks. For customisation by sound, each subject attended a ten to twenty-minute training session. Additional explanation was sometimes provided when needed.

5.8.3 Subjects

A total of 40 graduate and undergraduate students voluntarily participated, 20 each on small and large menu designs. These were split 15 / 5 and 16 / 4 respectively between males and females. We decided to have 20 subjects in each experiment, since the usability of any system can be sufficiently evaluated by seven to twenty users, according to Nielsen [137]. In addition, we felt that this number would provide us with vital data on the benefits and drawbacks of each approach, while keeping the experiment under control. The ages of subjects in both experiments ranged from 18 to 44 years, while their average computer usage exceeded 12 hours per week. In both experiments, each subject was randomly assigned to one of 4 groups of 5 subjects, each of which followed the three experimental menu conditions in a different order. Subjects were given one recorded tutorial according to the experiment they participated in and were asked to accomplish the same group of tasks (50 selections for each session in each condition).

5.9 Results

This section reports the experimental results in terms of both quantitative and qualitative measures, along with self-reported and observed data. In addition, interviews were conducted with subjects when needed. It was noticeable that participants preferred the positions of menu items not to be changed. On the other hand, subjects who participated in the evaluation of the adaptable large menu hesitated to customise it and before starting to use the speech recognition system, subjects lacked confidence and appeared not to want or to need to customise by sound.

5.9.1 Efficiency

5.9.1.1 Small Menu Selection Time

A one-way repeated measures ANOVA showed that there were no significant differences in efficiency among the three menu conditions overall at 0.05 ($F(2, 57) = 0.53$, $p < 0.59$). In addition, when the differences between the menu types were analysed using the t-test at 0.05, there were no significant differences in efficiency among the three menu conditions (Table 25 and Figure 42). This result indicates that for all three conditions the design drawbacks had been mitigated and the limitations reduced.

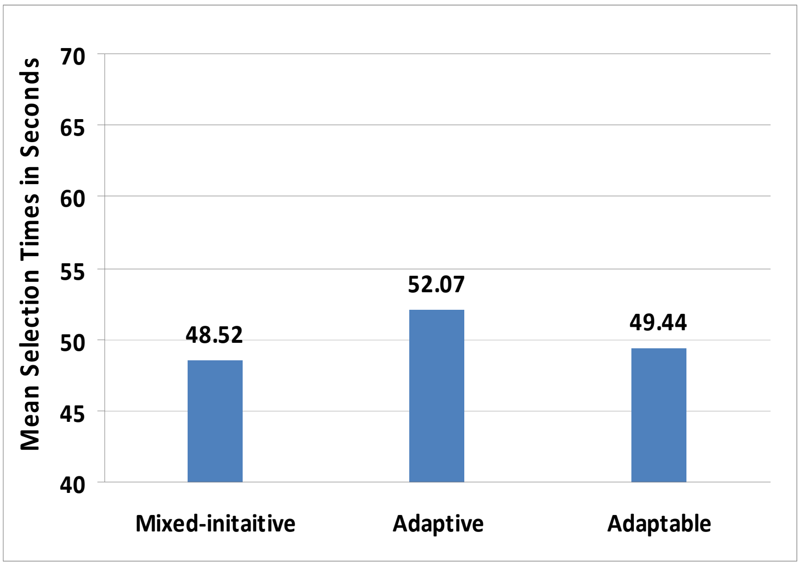


Figure 42: Mean selection times for small menus

Table 25: Results of t-test for main tasks

Conditions	Searching
Adaptable vs. Mixed-Initiative	t19 = 0.47, p = 0.65, r =0.1072
Adaptive vs. Mixed-Initiative	t19 = 1.90, p = 0.07, r =0.3996
Adaptable vs. Adaptive	t19 = 1.48, p = 1.16, r =0.3215

Note: Statistically significant results are displayed in **bold**

5.9.1.2 Large Menu Selection Time

A one-way repeated measures ANOVA showed in the case of large menus that there were significant overall differences in efficiency among the three conditions at 0.05 ($F(2, 54) = 0.77$, $p < 0.47$). When these differences were analysed using the t-test at 0.05, the results supported Hypothesis 8 (Section 5.7 above) that subjects were significantly faster when using the adaptive menu than the adaptable one ($t19 = 2.65$, $p = 0.016$, $r = 0.5195$), but not significantly faster compared to the mixed-initiative menu ($t19 = 0.61$, $p = 0.6$, $r = 0.1386$). In addition, subjects were significantly faster with the mixed-initiative menu than the adaptable menu ($t19=2.22$, $p < 0.05$, $r = 0.454$), supporting Hypothesis 9 (see Table 26 and Figure 43).

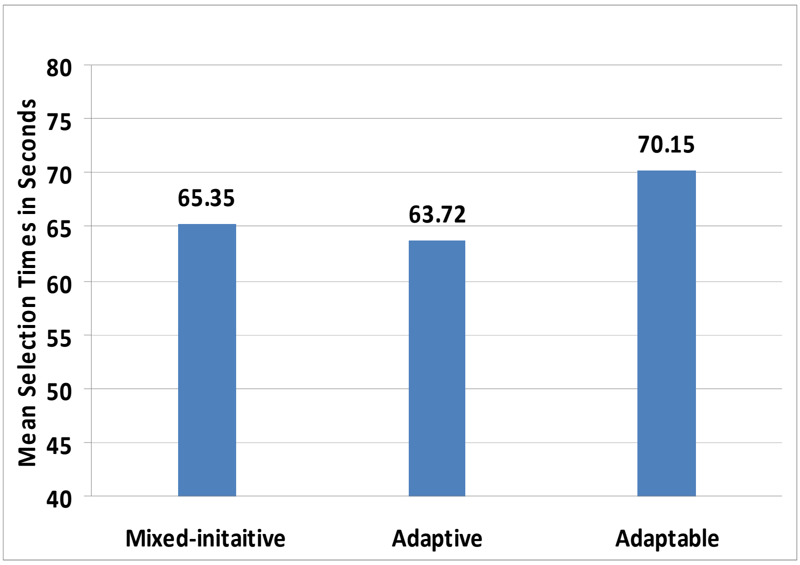


Figure 43: Mean selection times for large menus

Table 26: Results of t-test for main tasks

Conditions	Searching
Adaptable vs. Mixed-Initiative	**t19=2.22, p < 0.05, r =0.4538**
Adaptive vs. Mixed-Initiative	t19 = 0.61, p = 0.6, r =0.1386
Adaptable vs. Adaptive	**t19 = 2.65, p = 0.016, r =0.5195**

Note: Statistically significant results are displayed in **bold**

5.9.2 Error Rate

Table 27 shows the total number of errors for all subjects for each menu condition. An error was recorded when a subject clicked an item that was different from the target. Each cell contains the number of errors the subjects made in 1500 selections (50 selections x 1 session x 30 subjects) in each condition. It can be seen that the adaptable approach had the largest number of errors in both small and large menus, with 11 and 28 errors respectively, followed by the adaptive with 10 and 5 errors respectively. The mixed-initiative approach, as predicted in Hypothesis 9, had the fewest errors in both small and large menus, with 8 and 4 errors respectively. Surprisingly, the number of errors in the large mixed-initiative and adaptive menus were only half of those in their small counterparts.

Table 27: Frequency of user errors

Menu	Small	Large	Sum
Mixed-initiative	8	4	12
Adaptable	11	28	39
Adaptive	10	5	15
Total	29	37	66

5.9.3 Effectiveness

The effectiveness of the experimental environments was measured by calculating the percentage of tasks completed successfully by all subjects. A task would be regarded as successfully completed if and only if all the subjects who accomplished it using a particular environment completed it within its critical completion time. As Figure 44 shows, in small menus, 23.33% of tasks were successfully completed by all users using the adaptable menu and 20% using the mixed-initiative menu, whereas the rate for the adaptive menu was only 10%. By contrast, just 1 task was successfully completed by all users within the criterion time using the large adaptable and mixed-initiative menus.

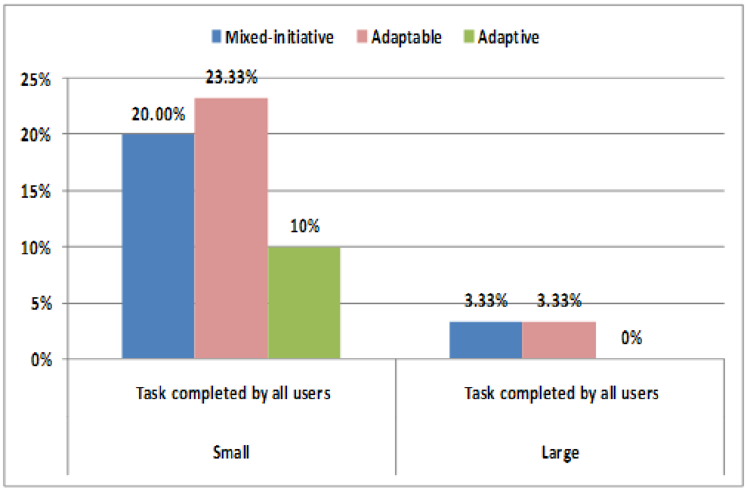


Figure 44: Effectiveness of small and large menus

In order to verify Hypotheses 8 to 11 stated in Section 5.7, the statistical t-test was used. Figure 44 shows that as expected, in small menus, the adaptable approach was more effective than the adaptive one in terms of number of tasks successfully completed. The t-test revealed that this difference was significant at 0.05 ($t_{19} = 1.71$, $p < 0.05$, $r = 0.3652$). Hypothesis 14 was also supported, as the mixed-initiative menu was more effective than the adaptive and the difference was significant at 0.05

Table 28: T-test for menu effectiveness

Conditions	Small menus	Large menus
	Tasks completed by all subjects	Tasks completed by all subjects
AV vs. MI	**$t_{19} = 1.37$, $p = 0.019$, $r = 0.2998$**	$t_{19} = 1.0$, $p = 0.32$, $r = 0.2236$
AD vs. MI	$t_{19} = 0.44$, $p = 0.066$, $r = 0.1004$	$t_{19} = 1.0$, $p = 0.32$, $r = 0.2236$
AD vs. AV	**$t_{19} = 1.71$, $p < 0.05$, $r = 0.3652$**	$t_{19} = 1.0$, $p = 0.32$, $r = 0.2236$

($t_{19} = 1.37$, $p = 0.019$, $r = 0.2998$). For large menus, however, the differences between the adaptable, adaptive and mixed-initiative conditions were found not to be statistically significant (see Table 28).

5.9.4 Customisation Behaviour

In the second experiment, subjects were given an opportunity to customise before starting the task, whereas in this third experiment, they were able to customise during the tasks using speech, but were not allowed to do so before clicking the items for the first time. This was done to ensure that subjects engaged with the tasks and were not distracted by customising the menus, as well as to ensure that they customised when they needed to, rather than because an item was included in the task. It was found that subjects spent significantly less time customising the small menus than the large ones, moving items 151 and 123 times respectively. In the small mixed-initiative menu, only two subjects failed to select the frequency technique, whereas three subjects did so in the large one.

5.9.4.1 Adaptation by users

Subject could easily move items up by saying the name of the required item. Subjects were trained how to customise and provided with help when needed (after repeating the name twice), since we were interested in the results of customisation, not the way in which it was done. It was observed that subjects utilised different criteria for ordering the menu items. The most common approach was frequency-based, but this did not prevent some subjects from using the alphabetical ordering approach. For example, one subject moved those items that started with the letter P. Subjects behaved differently towards adaptable menus according to their size; for example, they customised large menus more than small ones. It was also found that under the mixed-initiative condition, subjects utilised the frequency and recency techniques more in large than small menus, with respective totals of 112 and 105 selections made by subjects.

5.9.4.2 Adaptation by the system

The results of the second experiment on the adaptive menus (Section 4.7.1) showed that changing both the recently and frequently clicked items was more efficient than just changing or highlighting the frequently clicked items. In addition, subjects interviewed after the experiment commented that they had noticed that the mixed-initiative menu was split into three sections (frequently used, recently used and other items), whereas they thought that the adaptive split menu had only two sections (frequently used and other items).

5.9.5 User Satisfaction

At the end of the experiment subjects were asked to give ratings of 1 to 5 to indicate their preferences. As Figure 45 shows, the adaptable approach was the most strongly preferred in both large and small formats, with exactly half and almost half of users selecting it as their preferred approach respectively, whereas almost half of subjects selected the minimised approach as the least preferred in both formats.

Q: Which menu did you prefer overall? Put a 1 beside that menu and a 2 beside the next preferred menu.

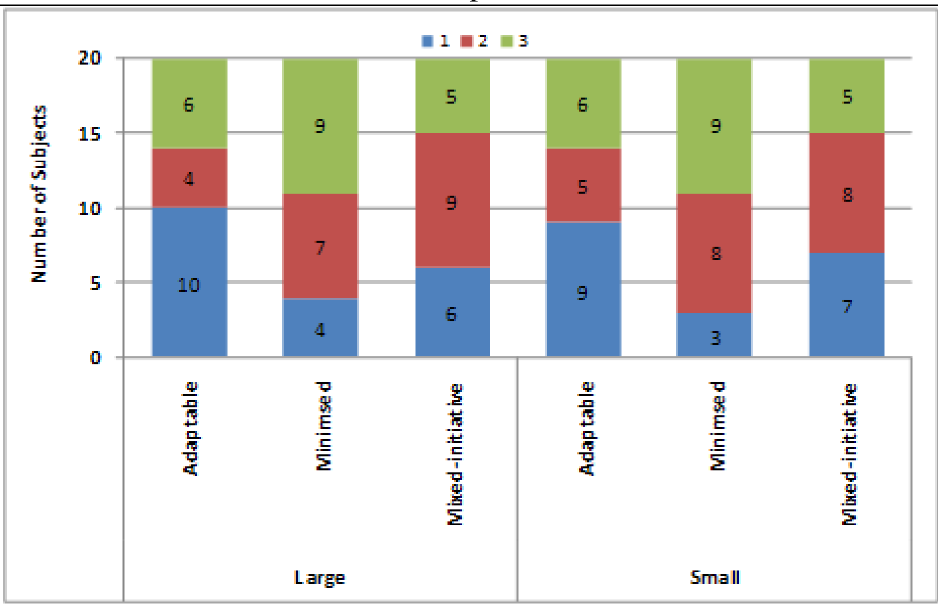


Figure 45: Preferences

Q: Which menu scheme did you feel the easiest to use? Put a 1 beside that menu and a 2 beside the next preferred menu.

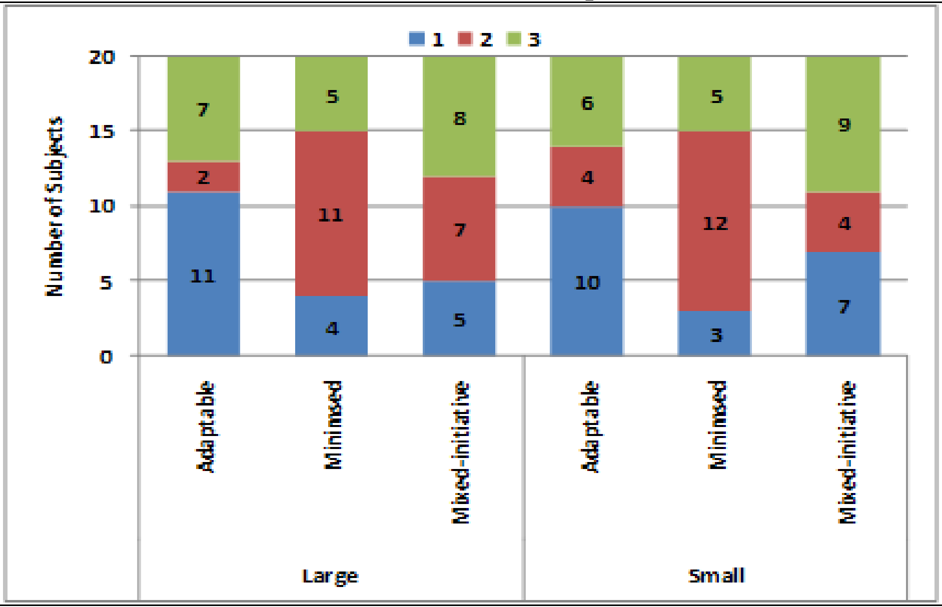


Figure 46: Ease of use

Overall, in support of Hypothesis 9, the mixed-initiative approach was the most strongly preferred, with exactly three-quarters of subjects rating it either first or second in both large and small menus. This was closely followed by the adaptable approach, which almost three-quarters of subjects placed either first or second in both formats. Figure 46 shows that the adaptable approach was perceived as the easiest to use in both large and small menus, with more than half and exactly half of subjects selecting it as the easiest approach respectively, whereas almost half of subjects selected the minimised approach as the least easy to use in both large and small menus (8 and 9 subjects respectively). Overall, the mixed-initiative menus were considered the easiest, with exactly three-quarters of subjects placing them either first or second in both large and small formats, closely followed by the adaptable approach, with almost two-thirds for the large menu and almost three-quarters for the small one. As shown in Figure 47, for small menus, 65% of subjects gave a score of 5 for sound as output and the remainder gave a score of 4. Scores for sound as input were lower: 40% of subjects rated this 5, 45% 4 and the rest 3. More than half (60%) of subjects gave a score of 5 for the proposition that sound matched their needs in small menus, 35% scored this 4 and the remaining 5% 3. In response to the final item, 45% of subjects gave small menus a score of 5 for ease of customising by sound, while half scored it 4 and the rest 3.

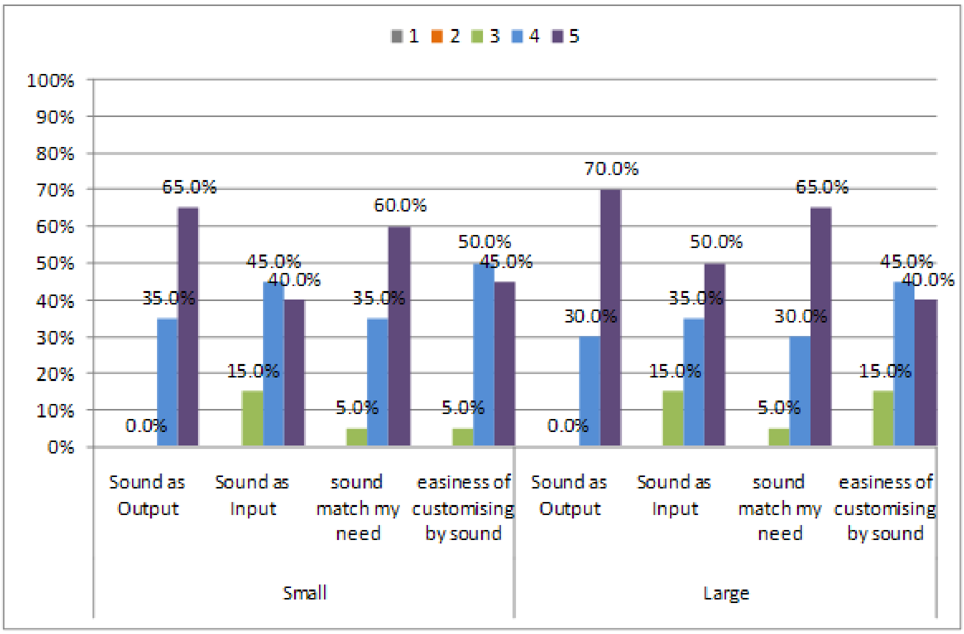


Figure 47: Subjects' opinions

The pattern was much the same for large menus: 70% of subjects scored 5 for sound as output and the rest 4; half of subjects rated sound as input at 5, 35% 4 and 15% 3; 65% scored 5 for the sound matching their needs, 30% 4 and the rest 3; and 40% of subjects scored 5 for ease of customising by sound, 45% 4 and the rest 3.

5.9.6 Customisation Behaviour

Subjects allowed had one opportunity to customise: after starting to use the adaptable menu and while users performing the tasks. In addition, they were not allowed to customise before clicking the menu items first. In other words, subjects were told to use the sound when they want to change the positions of the required items. Since, users do not know the positions of the items. Therefore, subjects not allowed to move items before clicking it first. The main reason behind that to ensure that users utilised sound when they needed. Figure 48 shows the percentage of subjects who customise the adaptable condition by sound and the percentage of helped required by users.

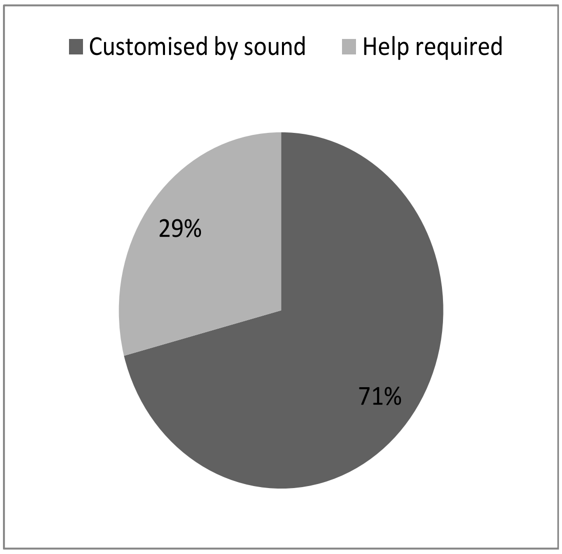


Figure 48: Percentage of customisation by sound and help required

5.10 Discussion

The aim of this chapter was to compare the three personalisation approaches: adaptive, adaptable and mixed-initiative. The design of the menus was unified (all were split menus) in order to eliminate the effect of utilising different designs. For small menus, the results show that there were no significant differences among the three approaches. A possible explanation for this is that the size of the menu helps subjects to remember the position of items. This is supported by the observation that some subjects preferred not to customise the adaptable menus. These results differ from those of previous studies, since improvements were made to the three menus in order to limit the design drawbacks. In large menus, the results show that subjects were significantly faster with the adaptive menu than the adaptable menu, but not significantly faster compared to the mixed-initiative type. In addition, subjects were significantly faster with the mixed-initiative menu than the adaptable menu. This finding is in agreement with that of Sears and Shneiderman [113], who report that subjects were faster using adaptive menus than adaptable ones. In addition, adaptable approach had the largest number of errors, with 11 and 28 errors respectively, followed by the adaptive with 10 and 5 errors respectively. The mixed-initiative approach, had the fewest errors in both small and large menus, with 8 and 4 errors respectively. Surprisingly, the number of errors in the large mixed-initiative and adaptive menus were only half of those in their small counterparts (see Figure 49).

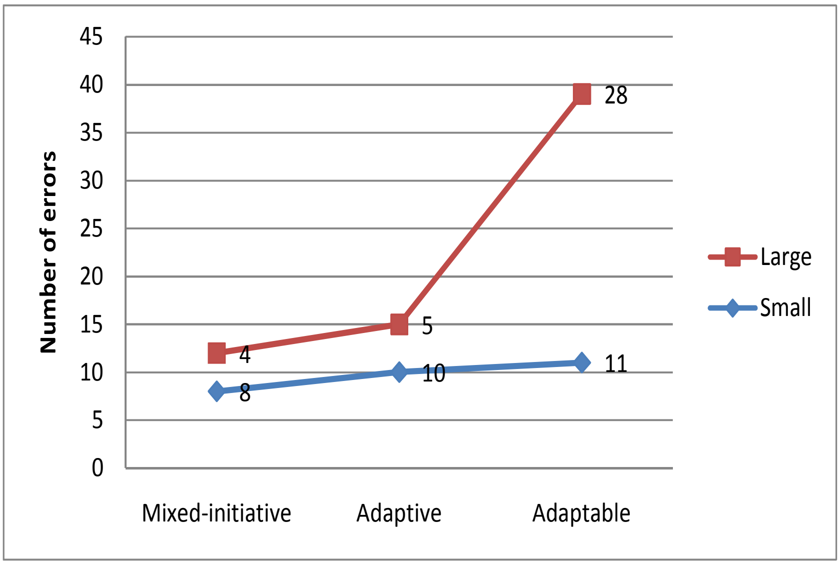


Figure 49: Frequency of user errors

For both small and large menus it was found that subjects preferred the mixed-initiative approach, followed by the adaptable approach. This suggests that subjects prefer to have full control at the same time as receiving some suggestions and assistance from the system. In addition, 15 subjects liked to customise by speech, whereas others found it difficult. Surprisingly, the number of errors in the large mixed-initiative and adaptive menus were only half of those in their small counterparts. In addition, the adaptable approach had the largest number of errors in both small and large menus.

An independent unpaired t-test was performed in order to determine the effect of using the ‘as you go’ customisation strategy, by comparing the customisation of the adaptable menu using speech with that of the same menu without using speech, conducted in session 1 of the second experiment (see Table 29). For small menus (Figure 50) the t-test revealed that the difference was not quite statistically significant at 0.05 ($t42 = 1.49$, $p = 0.074$, $r = 0.2241$).

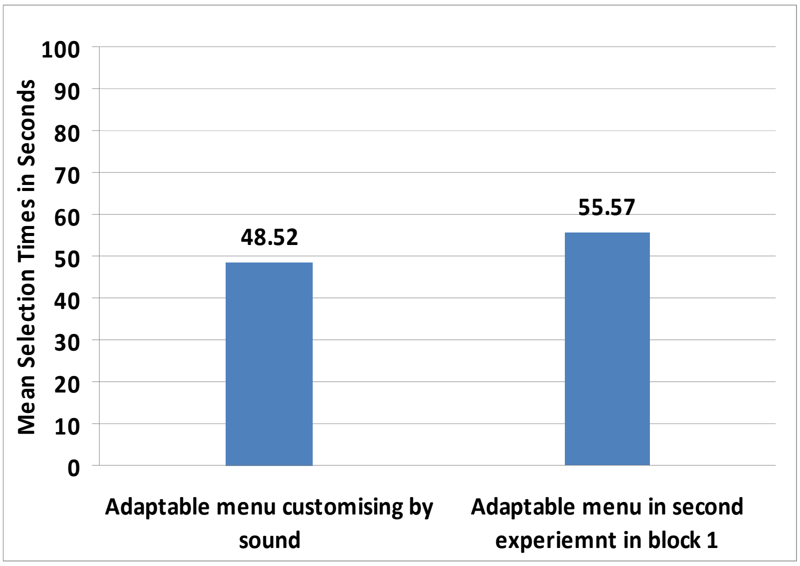


Figure 50: Selection time for small adaptable menus customised using sound vs. non-customised ones

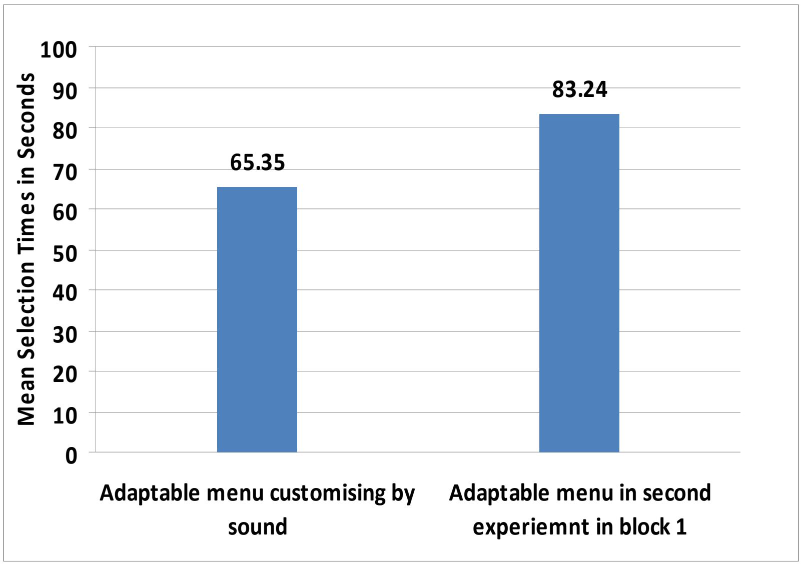


Figure 51: Selection time for large adaptable menus customised using sound vs. non-customised ones

For large menus the t-test revealed a statistically significant difference at 0.05 ($t_{44} = 1.99$, $p < 0.05$, $r = 0.2874$), supporting Hypothesis 6. It was found that selection time for large menus customised during the tasks using speech was faster than for non-customised menus, the average time for subjects who used speech being 65.35 seconds, compared with an average of 83.24 seconds for the fastest large menu in session 1 (Figure 51).

Table 29: Adaptable menus with vs. without speech

Adaptable with speech vs. adaptable in session 1	T-test
Small menus	($t_{42} = 1.49$, $p = 0.074$, $r = 0.2241$)
Large menus	**($t_{44} = 1.99$, $p = 0.021$, $r = 0.2874$)**

There was a variety of responses towards the design of each approach. First of all, in terms of design of the adaptive interface, subjects generally liked the way that the system assisted them by moving items to the top. However, there were comments suggesting that moving items continuously was confusing. In other words, there was a need for adaptation but with less movement. On the other hand, in terms of design of the adaptable interface, subjects generally liked the way of moving items by using sound. Furthermore, subjects were aware of the number of items displayed in each list. Last but not least, in terms of design of the mixed-initiative interface, subjects generally liked to be able to choose the techniques that they preferred. This technique made all subjects realise that there was a split into three sections. By contrast, in the adaptive menu, most subjects (14) thought that there were only two sections, not three; during interviews, many were surprised to be told that this type was also divided into three sections. This illustrates that in the adaptive approach subjects were not aware of the menu and its environment, whereas the mixed-initiative approach increased subjects' awareness, since they were involved as users in the customisation.

5.11 Summary

This chapter has documented the third set of experiments carried out to investigate the usability of three different types of personalised split menus. Usability was measured in terms of efficiency, effectiveness and satisfaction. Two metrics were used to measure efficiency: task accomplishment time and frequency of errors. Effectiveness was measured by the number of tasks completed successfully within task criterion time and satisfaction was measured on 5-point Likert scales.

The results show that in large menus, subjects were significantly faster with the adaptive menu than the adaptable menu, but not significantly faster compared to the mixed-initiative type. For both small and large menus it was found that subjects preferred the mixed-initiative approach, followed by the adaptable approach. Surprisingly, the number of errors in the large mixed-initiative and adaptive menus were only half of those in their small counterparts. In addition, the adaptable approach had the largest number of errors in both small and large menus. Furthermore, 15 subjects liked to customise by speech, whereas others found it difficult. From this result and others reported in Chapter 4, it can be concluded that size of content is one of the main factors affecting the usability of personalised approaches. Moreover, the usability of the three personalised approaches were found to differ according to the size of content. As the size of content increases, so the need for the adaptive and mixed-initiative approaches increases and the value of the adaptable approach decreases. Conversely, as the size of content decreases, so the value of the adaptable approach increases and that of the adaptive and mixed-initiative types decreases.

Chapter 6: Conclusions and Empirically Derived Guidelines

6.1 Introduction

The primary goal of this study was to examine the efficiency, effectiveness, user satisfaction and user customisation behaviour of three approaches to personalisation: adaptive, adaptable and mixed-initiative, since far too little attention has been paid to comparing the usability of these three approaches to either web content or graphical user interfaces. Additional goals were to mitigate the drawbacks of each approach and to examine which condition was then most usable. This chapter summarises the steps taken to achieve this goal, followed by an assessment of the contribution of this study, which elaborates on the findings from each step. Empirically derived guidelines and suggestions are then produced in order to elaborate the findings and produce more usable interfaces. The chapter concludes with recommendations for future work.

The first step in fulfilling the research aim was to examine the value of customisation in the three approaches to the personalisation of web content and graphical user interfaces. Therefore, Study One reported in Chapter 3 evaluated the usability of adaptive, adaptable and mixed-initiative approaches to web content, while Study Two did likewise for graphical user interfaces. The second step was to understand the factors causing one of these approaches to be successful at one time and unsuccessful at others. Studies Two and Three examined the effects of utilising different content sizes and applying different techniques to the adaptive, adaptable and mixed-initiative conditions.

6.2 Limitations

These studies were conducted as lab experiments and as with any such experiments there were limitations. For example, there are several variables which may affect users' ability to search for items and customise. These include working in a busy workplace, extended time between selections, working on different tasks at the same time, training, motivation and the provision of explanations of the benefits of customisation. In addition, our experiments were done in a short period of time and results might have been different over a longer period. Furthermore, these experiments tested web content and graphical user interfaces in very specific systems and environments. The question here is how to generalise these conditions to other environments. In addition, we utilised specific levels of adaptivity and adaptability, which suggests that the results might be different in the case of more sophisticated applications. Adaptivity and adaptability depend on the algorithms utilised, while adaptability also depends on ease of customisation. Since it would have been difficult to quantify the levels of adaptivity and adaptability, we simply provided the same level of these conditions throughout. In addition, regarding the customisation of menus using sound, questions remain as to how to generalise these conditions to other environments such as using a mobile phone in the street. Finally, there are limitations regarding simplicity: our experiments used a simple algorithm for adaptive and mixed-initiative conditions based on simple values of frequency and recency; in addition, we used a simple customisation. It may be that more sophisticated algorithms would behave differently.

6.3 Addressing the Overall Hypothesis

This section reviews the acceptance and rejection of overall research hypotheses. Table 30 reviews the acceptance and rejection status with the probability level of the overall research hypothesis according to its five parts and ten sub-hypotheses.

Table 30: Review of overall hypothesis acceptance and rejection.

	Hypothesis	**Accepted**	**Rejected**	**Reference**
H1	The adaptive, adaptable and mixed-initiative approaches will be more efficient than the static approach in terms of:			Section 3.5
	▪ Task accomplishment time.	√		
	▪ Frequency of clicks and pages visited.	√		
	▪ Number of tasks completed successfully.	√		
	▪ Preferences.	√		
H2	The adaptive approach will be more efficient than the adaptable approach in terms of:			Section 3.5
	▪ Task accomplishment time.		√	
	▪ Frequency of clicks and pages visited.		√	
	▪ Number of tasks completed successfully.		√	

	▪ Preferences.		√	
H3	The mixed-initiative approach will be more efficient than both the adaptive and adaptable approaches in:			Section 3.5
	▪ Task accomplishment time.	√		
	▪ Frequency of clicks and pages visited.	√		
	▪ Number of tasks completed successfully.	√		
	▪ Preferences.	√		
H4	The adaptable and mixed-initiative approaches will be more adequate than both the static and adaptive approaches in terms of:			Section 3.5
	▪ Controllability.	√		
H5	In small menus, the adaptable approach will be more efficient than the adaptive one in terms of:			Section 4.5
	▪ Task accomplishment time.	√		
	▪ Frequency of errors.	√		
	▪ Number of tasks completed successfully.	√		
	▪ Preferences.	√		
H6, H8	In large menus, the adaptive approach will be more efficient than the adaptable one in terms of:			Rejected on section 4.5 and accepted on 5.7
	▪ Task accomplishment time.	√	√	
	▪ Frequency of errors.	√	√	
	▪ Number of tasks completed successfully.	√	√	
	▪ Preferences.	√	√	
H7, H9	In small and large menus, the mixed-initiative approach will be more efficient than both the adaptive and adaptable conditions in terms of:			Rejected on section 4.5 and accepted on 5.7
	▪ Task accomplishment time.	√	√	
	▪ Frequency of errors.	√	√	
	▪ Number of tasks completed successfully.	√	√	
	▪ Preferences.	√	√	
H10	In small menus, the adaptable menu customised by sound will be more efficient than the same menu not customised by sound in terms of:			Section 5.7
	▪ Task accomplishment time.	√		
H11	In large menus, the adaptable menu customised by sound will be more efficient than the same menu not customised by sound in terms of:			Section 5.7
	▪ Task accomplishment time.	√		

6.4 Achievement of Research Goals

6.4.1 Comparison of Approaches to Personalisation of Web Contents

Study One was designed to address the primary aim of comparing the usability of static, adaptive, adaptable and mixed-initiative approaches to personalisation of web content. The results show that subjects performed fastest in the mixed-initiative condition, which they also preferred overall to the other approaches.

6.4.2 Comparison of Approaches to the Personalisation of Interfaces

Study Two was designed to address the second aim of comparing the usability of adaptive, adaptable and mixed-initiative approaches to the personalisation of graphical user interfaces. Addressing the hypotheses posed at the beginning of this study and the findings, it is now possible to state the following conclusions:

For efficiency, we conclude that in small menus, the adaptable approach was surprisingly the best in term of efficiency, whereas in large menus, the split menu condition was found to be the best on initial use and the minimised on the second time of use (see section 4.7.1).

For frequency of error-occurrence, we conclude that in small menus, errors were reduced in the adaptable menu by 50% when it was customised by subjects, whereas they were increased by 50% in large ones. In addition, errors decreased in large menus by 50% when menus were customised by the system (see section 4.7.2).

For effectiveness, we conclude that in small menus, both the number of subjects who completed all tasks and the number of tasks completed by all subjects was higher under the adaptable condition than any other condition. On the other hand, for large menus alone, the number of subjects who completed all tasks was highest for the split condition and the number of tasks completed was highest for the highlighted condition, while fewer subjects completed all tasks using the minimised menu than other types and the number of tasks completed by all subjects was lowest under the adaptable and mixed-initiative conditions. This leads us to conclude that in small menus the adaptable approach was more effective than other conditions, whereas in large menus the adaptive approach, such as highlighted and split menus, was more effective than other conditions (see section 4.7.3).

For satisfaction, we conclude that in small menus, the minimised type was the most strongly preferred, as more than half of subjects ranked it either first or second. This was followed by the adaptable and highlighted menus with exactly half and more than one-third of subjects ranking them first or second respectively. The least strongly preferred types (by 17 subjects each) were the split and mixed-initiative menus. However, 11 of the 17 subjects ranked the split menu last, while 8 did so for the mixed-initiative type, suggesting that the former was the less preferred of the two. On the other hand, for large menus, the mixed-initiative type was ranked first by thirteen subjects, with more than two-thirds ranking it either first or second, followed by the minimised menu, ranked first or second by more than half of subjects. The least desirable was the adaptable type, followed by the split menu (see section 4.7.4).

6.4.3 Comparison of Improved Approaches to Personalisation

Study Three was designed to compare the usability of improved adaptive, adaptable and mixed-initiative approaches to the personalisation of graphical user interfaces. Returning to the hypotheses posed at the beginning of this study and the findings, it is now possible to state the following conclusions:

For efficiency, we conclude that in small menus, after mitigating the design drawbacks and disadvantages of personalised approaches, there was no significant difference in efficiency among the three menu conditions. In large menus, subjects were significantly faster when utilising the adaptive and mixed-initiative conditions than the adaptable one (see section 5.9.1).

For frequency of error-occurrence, we conclude that in small menus, errors were reduced in the mixed-initiative and adaptive conditions by 50%, while they increased by 50% in large menus when these were customised by subjects (see section 5.9.2).

For effectiveness, we conclude that in small menus, almost a quarter of tasks were successfully completed by all users within the criterion time using the adaptable menu and 20% by using the mixed-initiative type, whereas the number of tasks successfully completed by all users within the criterion time using the adaptive approach was only 10%. On the other hand, in large menus, just one task was successfully completed by all users within the criterion time using the adaptable and mixed-initiative menus (see section 5.9.3).

For satisfaction, we conclude that the mixed-initiative menu type was the most strongly preferred, with exactly three-quarters of subjects choosing it as either their first or second preference in both large and small menus. This was closely followed by the adaptable approach, which almost three-quarters of subjects made their first or second preference in both large and small menus (see section 5.9.5).

6.5 Empirically Derived Guidelines

Considering again the hypotheses posed at the beginning of this thesis, it is now possible to offer empirically derived guidelines intended to increase the usability of personalised content and menus and to enhance the interaction between users and applications. This section derives a number of usability guidelines that can used to design more usable personalised interfaces. These are presented in order to assist designers and individual users in deciding which of static, adaptive, adaptable and mixed-initiative approaches should be used in personalising web content and graphical user interfaces. This is done by considering what level of customisation should be provided, what techniques should be used and the appropriate environment. The following guidelines are linked to the results obtained, which helps us to draw conclusions and which should provide more helpful advice to designers. These empirical guidelines are presented in four subcategories.

6.5.1 Designing a Personalised Interface

The experiments in this thesis suggest that designers need to consider the different characteristics of personalisation approaches. They must consider the size of content and the complexity of graphical user interfaces. In addition, they must consider the level of personalisation provided to users. This involves the use of guidelines on size of content, interface complexity and level of customisation.

6.5.1.1 Size of Content and Interface Complexity

Before choosing one of the personalisation approaches, designers must consider the size of content and the complexity of the interface, since our results show that adaptive, adaptable and mixed-initiative conditions behave differently based on the size of content, which makes the graphical user interface more complex (see section 5.9). These differences were explored in Study Two. Therefore, designers should consider the following:

As the size of content and the visual complexity of the interface increase, the need for the adaptive and mixed-initiative approaches grows and the suitability of the adaptable approach decreases, and vice-versa. This is proposed because of the result reported in section 5.9.1, supporting Hypothesis 3, stated in Section 5.7, that in large menus subjects were significantly faster when utilising adaptive menus than adaptable ones ($t19 = 2.65$, $p = 0.016$, $r = 0.07$), but not significantly faster compared to the mixed-initiative ($t19 = 0.61$, $p = 0.6$, $r = 0.03$). For example, during these sets of experiments it was noticeable that once the size of content became large and the graphical user interface became complex users welcomed suggestions. For instance, in the mixed-initiative menu, users ignored suggestion made in small menus but welcomed them in large ones.

6.5.1.2 Adaptability Usage

One of the main objectives of this study was to investigate the effect of different levels of adaptability. Therefore, we conducted in experiment two (see section 4.7.1) session 2 a comparison of three adaptable menus presented with different types of adaptability: (1) help not provided (that is, adaptable menu), (2) assistance provided by highlighting the frequently clicked items (that is, highlighted menu) and (3) recommendations provided by moving frequently clicked items to the top of the list, followed by a horizontal line separating the recently clicked items and hiding the others (that is, minimised menu). Our findings show an effect of personalisation level on users' performance towards level of content (see section 4.7.1). However, there are a number of issues raised by the experiment which a designer must take into consideration when using different levels of adaptability.

1. As the size of content and visual complexity of the interface increases, a higher level of adaptability is needed and vice-versa. This is proposed because the adaptable approach (low) was found to be the best for small menus and the minimised (high) for large ones (see section 4.7.1). For example, during these sets of experiments it was noticeable that users liked to feel in control all the time but once the visual complexity of the interface increased they welcomed system recommendations and interruptions.

2. Users prefer and trust recommendations provided by the system. This is shown by the fact that no subjects changed or examined suggestions from the system in

minimised menus. For example, in the second experiment when the subjects attempted to customise the minimised menu, a new recommendation was provided by ordering the menu items based on usage. None of the participants examined the order of the menu; all trusted the menu recommendations and believed that it was the best order without giving it more than a quick glance.

6.5.1.3 Adaptivity Usage

A second objective was to investigate the effects of different levels of adaptation. Therefore, we conducted a comparison of three levels of adaptive menus in experiment two (see section 4.7.1). In session 1, participants were presented with different types of adaptation: (1) changes occurring without moving items (that is, highlighted menu), (2) changes made by moving recently and frequently clicked items to the top of the list and leaving the others unchanged (that is, split menu) and (3) changes made by moving only frequently clicked items to the top of the list and leaving the others unchanged (that is, minimised menu). There was an issue arising from the experiment which designers should take into consideration when using different levels of adaptivity.

As the size of content and the visual complexity of interfaces increase, so the need for a high level of adaptivity increases and vice-versa. This is proposed because the t-test revealed that in small menus subjects were significantly faster with the highlighted menu (low) than the minimised menu ($t_{29} = 2.8$, $p < 0.01$, $r = 0.46$) and the split menu (high) ($t_{29} = 2.17$, $p < 0.05$, $r = 0.38$), whereas in large menus subjects were significantly faster with the adaptive split (high) than the other conditions.

6.5.1.4 Mixed-initiative Usage

Direct comparisons between mixed-initiative and adaptable or adaptive menus are rare. In our evaluation of mixed-initiative menus we found that system recommendations tended to improve efficiency, effectiveness and user satisfaction. In all experiments conducted as part of this study, the mixed-initiative approach was preferred to the adaptive and adaptable conditions. The findings of this study suggest that further experimental investigation of the mixed-initiative approach is strongly recommended. Meanwhile, there are a number of issues arising from the experiments conducted here which designers should take into consideration when following a mixed-initiative approach.

1. They must ensure that users are aware at all times of the suggestions made by the system, since these may occur while users are busy. This is proposed because it was noticeable that in the second experiment subjects were not aware of menu recommendations. To solve this problem we utilised earcons to indicate that suggestions had been made by the system. For example, in the second experiment while users were performing their tasks using the mixed-initiative menu, the system recommended the new frequency technique to allow users to change the menu items based on the frequency of selected items, but users were not aware of this recommendation.

2. They should ensure that users can overrule the recommendations made by the system. This is proposed because it was noticed that in the first experiment subjects liked being able to lock and unlock the lists of items. For example, in the first experiment users commented positively when they overruled the changes made by the list.

6.5.1.5 Initiative Usage

In the first and second experiments, the system waited for the user to initiate customisation before providing any support, except in the mixed-initiative menus, where the system initiated the making of suggestions to users. In addition, it was observed in the second and third experiments that subjects liked the fact that the system took the initiative by providing suggestions regarding the recency technique before the menu started to be used and by providing further suggestions once there were some frequently selected items. Therefore, designers should take into consideration the following issue arising from these experiments when taking the initiative.

Users preferred customisation suggestions made by the system, because customisation is a heavy task that most users attempt to avoid. In our experiment users welcomed customisation suggestions from the system. For example, participants liked the recency and frequency suggestions made for them in the mixed-initiative menus.

6.5.1.6 The Use of Speech

Speech was utilised in the third experiment in order for users to customise their menus as an incremental customisation strategy while they were performing their tasks. Such incremental customisation was found to be a powerful method to increase performance. A particular issue arising from the experiment which designers should take into consideration when using speech is the following.

We recommend that designers should not neglect the use of the auditory channel and speech to customise graphical user interfaces. The value of this recommendation increases if the visual channel is busy and if the content is large or the interface is visually complex. This is proposed because in large menus the t-test revealed a difference considered to be statistically significant at 0.05 ($t44 = 1.99$, $p < 0.05$, $r =$), as predicted in Hypothesis 6, stated in Section 5.7. This finding reveals that customising menus during the tasks by using speech is faster than not doing so. For example, in the third experiment subjects hesitated to use speech to move items to the top of the list, but once they had tried customising by sound they started to like it. In addition, it was noticeable that after a few seconds users started to become familiar with customising by sound while performing their tasks. One user commented that using speech in customisation was "a great idea to change by voice while I am busy doing something else".

6.5.1.7 Non-Personalised Techniques

It was found during interviews with users that they preferred to use personalised and non-personalised techniques, such as sorting (e.g. alphabetical order) and visualisation (e.g. grouping icons), since they depended heavily on the visual channel.

We therefore suggest that designers should mix personalised and non-personalised techniques. This is proposed because users suggested the use of colour to indicate frequently and recently selected menu items. In addition, some suggested that it would be helpful to sort the recently and frequently selected items by alphabetical order, whereas other users suggested doing so according to the length of the words.

6.6 Lessons Learnt

Throughout the time spent on this study, many lessons were learnt and the most important are listed here.

6.6.1 Making knowledge work

In order to discover new ideas and make them work, researchers need to investigate by reading, asking questions and writing notes on the work of other researchers. I learnt that to discover new ideas researchers need to open their minds to other people's work (especially unrelated to their own work). This gives the researcher a broader view of the world around him. Our world is full of details researchers should look at these from different perspectives and angles.

6.6.2 Pilot Test

Pilot testing is a good method to save time and effort in research by examining the proposed hypotheses and reducing the number of problems. The main benefit of a pilot test is the chance it gives the researcher to examine new ideas, approaches and clues. Therefore, I learnt that conducting a pilot test is required to observe carefully both the system and users' behaviour and then to note it. In addition, a pilot test gives brief and quick feedback about the environment and how users would normally react while using the systems being examined. This will help to gain good feedback from users about these systems. In addition, it was noticeable that some users did not know how to run the experiment but they naturally knew if something unusual happened or went wrong. Therefore, feedback from users is very important and researchers need to look at it from their perspective. In addition, I learnt that the chosen sample in the pilot test should be homogenous because this will help to get different feedback from different perspectives. For example, expert users can propose advanced ideas to improve the system's performance, while naïve users will propose advice on ideas to improve its ease of use.

6.6.3 Experiments Consistency

There are different methods, designs and approaches that can be utilised to conduct an experiment. The analysis of the results can also be tested statistically in a number of ways. Therefore, these methods should be decided at an early stage to ensure experimental consistency. Furthermore, it is very important to control the experiment to limit any effects on the experiment goals. For example, any independent variables need to be controlled during the experiment to ensure consistency. However, selecting the correct number of users is a critical issue in the research field, since opinions vary on this. For example, some researchers argue that 8 to 12 users is ideal, while others say that it should be more than that. However, I have learnt that keeping the experiment under control and making it easy to manage is very important and it helps to ensure consistency. In addition, I found that comparing two systems required the tasks and content of these two conditions to be equal to ensure experimental consistency

6.7 Directions for Future Research

6.7.1 Comparison of Personalisation for Different User Groups

Approaches to the personalisation of both web content and graphical user interfaces need to be explored with different groups of users, since our sample was quite homogeneous. Examples of such groups would be expert users, naïve users, children and the elderly, since some of these are likely to adapt easily to adaptive or adaptable menus, while others may find them difficult.

6.7.2 First Use vs. Frequent Use

Personalisation approaches to both web content and graphical user interfaces need to be explored by differentiating between first-time use and frequent use, since our experiments examined first-time use only and our results showed that using one approach, such as adaptive, for frequent use might be difficult for some users. In addition, using the adaptable approach for frequent use might make users more familiar with customisation and thus affect the results.

6.7.3 Other Menu Types

We examined different types of menu – highlighted, split and minimised – in the second experiment and split menus in the third. Menus of these types were presented as full sets, where all items appeared. There would be value in exploring other types of menu such as those like Microsoft personalised menus, which display a short set of items while others are hidden. It would also be useful to examine menus with subsets, which display a full set and another subset as a tree.

6.7.4 Other Sizes

Personalisation approaches to both web content and graphical user interfaces need to be explored with different screen sizes, such as mobile phones, since our experiments were limited to 17-inch monitors. The performance differences among the four approaches may in fact be larger when displaying large content on small screens.

6.7.5 Starting the Initiative

In mixed-initiative approaches, control is shared between the system and the user. Evaluation of this interaction has received little attention in the research literature. One could examine when and how much initiative the system should take to provide users with support for customisation. This support is helpful for users who forget to customise, but it might not be for users who do not want to do so. Therefore, the balance of initiative between users and the system should be carefully examined. It may be, for example, that users find systems which take more initiative to be annoying. In addition, it would be valuable to explore how much users really need suggestions while they are performing their tasks.

6.7.6 Comparing Personalisation Approaches to Alternative Techniques

6.7.6.1 Multimedia

We compared static, adaptive, adaptable and mixed-initiative approaches to personalisation of text in both graphical user interfaces and web content. One could also examine these approaches by combining different media such as text and graphics, text, graphics and speech, earcons and auditory icons. Such multimedia combinations might facilitate tasks or introduce new difficulties. For example, combining text with graphics might make searching less difficult than with text only.

6.7.6.2 Visual Complexity

Personalisation approaches tested in this study should be retested with different levels of visual complexity. For example, one could make comparisons between static, adaptive, adaptable and mixed-initiative approaches with full-feature interfaces on one hand and initial ones on the other, since working with full-feature interfaces is different from using initial interfaces. Our results show that size affects the usability of personalisation approaches; therefore, working in full-feature mode may make the adaptive approach relatively more difficult than the adaptable approach.

6.8 Epilogue

The empirical work reported in this thesis has demonstrated that the usability of customisation approaches is affected by a number or factors, one of the main ones being size of content: as this increases, the need of the adaptive and mixed-initiative approaches increases and that of the adaptable approach decreases. Conversely, as the size of content decreases, so the need of the adaptable approach increases and that of the adaptive and mixed-initiative ones decreases. It has also been shown that speech can be effectively utilised as an input metaphor in an incremental manner to customise graphical user interfaces and that it can effectively enhance the usability of interfaces and content as an output metaphor.

The guidelines presented in this thesis and the experimental findings will provide some help in understanding how to customise in more usable ways. However, more experimentation and investigation are also required to further improve our understanding of personalisation.

References

[1] J. Park, S. H. Han, Y. S. Park, and Y. Cho, "Adaptable versus adaptive menus on the desktop: Performance and user satisfaction," *International Journal of Industrial Ergonomics*, vol. 37, pp. 675-684, 2007.

[2] L. Kaufman and B. Weed, "Too much of a good thing?: identifying and resolving bloat in the user interface," *ACM SIGCHI Bulletin*, vol. 30, pp. 46-47, 1998.

[3] J. McGrenere and G. Moore, "Are We All In the Same" Bloat"?," 2000.

[4] K. Al-Omar and D. Rigas, "Comparison of Adaptive, Adaptable and Mixed-initiative Menus," presented at Cyberworlds Bradford, UK, 2009, in press.

[5] D. S. Lee and W. C. Yoon, "Quantitative results assessing design issues of selection-supportive menus," *International Journal of Industrial Ergonomics*, vol. 33, pp. 41-52, 2004.

[6] J. E. McDonald, M. E. Molander, and R. W. Noel, "Color-coding categories in menus," 1988.

[7] L. Findlater and J. McGrenere, "A comparison of static, adaptive, and adaptable menus," *Proceedings of the 2004 conference on Human factors in computing systems*, pp. 89-96, 2004.

[8] B. Shneiderman and P. Maes, "Direct manipulation vs. interface agents," *interactions*, vol. 4, pp. 42-61, 1997.

[9] J. McGrenere, R. M. Baecker, and K. S. Booth, "An evaluation of a multiple interface design solution for bloated software," *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, pp. 164-170, 2002.

[10] K. Höök, "Steps to take before intelligent user interfaces become real," *Interacting with computers*, vol. 12, pp. 409-426, 2000.

[11] T. Tsandilas, "User-controlled link adaptation," 2003.

[12] J. Fink, A. Kobsa, and A. Nill, "Adaptable and adaptive information provision for all users, including disabled and elderly people," *The New Review of Hypermedia and Multimedia*, vol. 4, pp. 163-188, 1998.

[13] C. Stephanidis, A. Paramythis, M. Sfyrakis, A. Stergiou, N. Maou, A. Leventis, G. Paparoulis, and C. Karagiannidis, "Adaptable and Adaptive User Interfaces for Disabled Users in the AVANTI Project," *Proceedings of the 5th International Conference on Intelligence and Services in Networks: Technology for Ubiquitous Telecom Services*, pp. 153-166, 1998.

[14] D. Billsus, C. A. Brunk, C. Evans, B. Gladish, and M. Pazzani, "Adaptive interfaces for ubiquitous web access," *Communications of the ACM*, vol. 45, pp. 34-38, 2002.

[15] G. J. F. Jones and P. J. Brown, "Context-Aware Retrieval for Ubiquitous Computing Environments."

[16] K. Gajos, D. Christianson, R. Hoffmann, T. Shaked, K. Henning, J. J. Long, and D. S. Weld, "Fast and robust interface generation for ubiquitous applications," *Proc. of Ubicomp*, pp. 37-55, 2005.

[17] L. Findlater and J. McGrenere, "Impact of screen size on performance, awareness, and user satisfaction with adaptive graphical user interfaces," 2008.

[18] D. Benyon, "Adaptive systems: a solution to usability problems," *User Modeling and User-Adapted Interaction*, vol. 3, pp. 65-87, 1993.

[19] P. W. Jordan, *An introduction to usability*: CRC, 1998.

[20] A. Abran, A. Khelifi, W. Suryn, and A. Seffah, "Usability Meanings and Interpretations in ISO Standards," *Software Quality Journal*, vol. 11, pp. 325-338, 2003.
[21] R. Oppermann, "Adaptively supported adaptability," *International Journal of Human-Computer Studies*, vol. 40, pp. 455-472, 1994.
[22] A. Dickinson, R. Eisma, and P. Gregor, "Challenging interfaces/redesigning users," *ACM SIGCAPH Computers and the Physically Handicapped*, pp. 61-68, 2002.
[23] D. I. Rigas, "Guidelines for auditory interface design An empirical investigation," 1996.
[24] S. A. Brewster, "The design of a sonically-enhanced interface toolkit," *Department of Computing Science Technical Report TR-1996-23, June*, 1996.
[25] T. Lindberg and R. Näsänen, "The effect of icon spacing and size on the speed of icon processing in the human visual system," *Displays*, vol. 24, pp. 111-120, 2003.
[26] S. A. Brewster, "Using non-speech sound to overcome information overload," *Displays*, vol. 17, pp. 179-189, 1997.
[27] M. L. Brown, S. L. Newsome, and E. P. Glinert, "An experiment into the use of auditory cues to reduce visual workload," *ACM SIGCHI Bulletin*, vol. 20, pp. 339-346, 1989.
[28] S. Brewster, "Using earcons to improve the usability of a graphics package," *PEOPLE AND COMPUTERS*, pp. 287-302, 1998.
[29] D. Rigas, H. Yu, K. Klearhou, and S. Mistry, "Designing Information Systems with Audio-Visual Synergy: Empirical Results of Browsing E-Mail Data," 2001.
[30] S. A. Brewster, P. C. Wright, A. J. Dix, and A. D. N. Edwards, "The sonic enhancement of graphical buttons," 1995.
[31] A. J. Dix and S. A. Brewster, "Causing trouble with buttons."
[32] S. A. Brewster and C. V. Clarke, "The design and evaluation of a sonically enhanced tool palette," *ACM Transactions on Applied Perception (TAP)*, vol. 2, pp. 455-461, 2005.
[33] S. A. Brewster and M. G. Crease, "Making menus musical," *Proceedings of IFIP Interact'97*, pp. 389-396, 1997.
[34] S. A. Brewster and M. G. Crease, "Correcting menu usability problems with sound," *Behaviour & Information Technology*, vol. 18, pp. 165-177, 1999.
[35] R. Direct, "Customization: Giving Users What They Want," Association Technology Forum & Idea Exchange, 2003.
[36] C. Stephanidis and M. Sfyrakis, "4.9. Current Trends in Man-Machine Interfaces: Potential Impact on People with Special Needs."
[37] A. Jaimes and N. Sebe, "Multimodal human–computer interaction: A survey," *Computer Vision and Image Understanding*, vol. 108, pp. 116-134, 2007.
[38] L. Chen and P. Pu, "Hybrid critiquing-based recommender systems," 2007.
[39] A. Tiwari, R. A. Hosn, S. H. Maes, I. Center, and N. Y. Yorktown Heights, "Conversational multi-modal browser: an integrated multi-modal browser and dialog manager," 2003.
[40] F. Paterno and C. Mancini, "Designing web interfaces adaptable to different types of use," 1999.
[41] L. Nigay and J. Coutaz, "A design space for multimodal systems: concurrent processing and data fusion," 1993.

[42] M. Jovanovic, D. Starcevic, and Z. Obrenovic, "Designing Aircraft Cockpit Displays: Borrowing from Multimodal User Interfaces," *Transactions on Computational Science III*, pp. 55, 2009.
[43] I. Oakley, M. R. McGee, S. Brewster, and P. Gray, "Putting the Feel in 'Look and Feel'."
[44] M. Beaudouin-Lafon and S. Conversy, "Auditory illusions for audio feedback," 1996.
[45] C. S. Campbell, S. Zhai, K. W. May, and P. P. Maglio, "What you feel must be what you see: adding tactile feedback to the trackpoint," 1999.
[46] W. W. Gaver, "Auditory Icons: Using Sound in Computer Interfaces," *Human-Computer Interaction*, vol. 2, pp. 34, 1986.
[47] W. Buxton, "Introduction to this special issue on nonspeech audio," *Human Computer Interaction*, vol. 4, pp. 1-9, 1989.
[48] W. W. Gaver, R. B. Smith, and T. O'Shea, "Effective sounds in complex systems: The ARKola simulation," 1991.
[49] J. L. Alty, "Can we use music in computer-human communication?," 1995.
[50] J. Sodnik, C. Dicke, S. Tomaži, and M. Billinghurst, "A user study of auditory versus visual interfaces for use while driving," *International Journal of Human-Computer Studies*, vol. 66, pp. 318-332, 2008.
[51] Ž. Obrenovic, D. Starcevic, and E. Jovanov, "Multimodal Presentation of Biomedical Data."
[52] W. W. Gaver, "The SonicFinder: An interface that uses auditory icons," *Human-Computer Interaction*, vol. 4, pp. 67-94, 1989.
[53] R. Lim, D. R. Perrott, and J. A. Gallegos, "Aurally aided visual search in depth," *The Journal of the Acoustical Society of America*, vol. 102, pp. 3139, 1997.
[54] D. R. Perrott, T. Sadralodabai, K. Saberi, and T. Z. Strybel, "Aurally aided visual search in the central visual field: effects of visual load and visual enhancement of the target," *Human Factors*, vol. 33, pp. 389, 1991.
[55] B. N. Walker, A. Nance, and J. Lindsay, "SPEARCONS: SPEECH-BASED EARCONS IMPROVE NAVIGATION PERFORMANCE IN AUDITORY MENUS."
[56] C. Nicol, S. Brewster, and P. Gray, "Designing sound: Towards a system for designing audio interfaces using timbre spaces," 2004.
[57] D. K. Mc Gookin, "The presentation of multiple earcons in a spatialised audio space," *Proceedings of BCS-HCI 2002*, vol. 2, pp. 228–229, 2002.
[58] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction*, vol. 4, pp. 11-44, 1989.
[59] D. A. Sumikawa, M. M. Blattner, K. I. Joy, and R. M. Greenberg, "Guidelines for the syntactic design of audio cues in computer interfaces," 1986.
[60] S. A. Brewster, "Providing a structured method for integrating non-speech audio into human-computer interfaces," *Department of Computer Science*, 1994.
[61] S. Brewster, A. Capriotti, and C. Hall, "Using compound earcons to represent hierarchies," *HCI LETTERS*, vol. 1, pp. 6-8, 1998.
[62] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "A detailed investigation into the effectiveness of earcons," 1994.
[63] G. Leplâtre and S. A. Brewster, "Designing non-speech sounds to support navigation in mobile phone menus," 2000.

[64] S. Brewster, P. C. Wright, and A. D. N. Edwards, "Parallel earcons: Reducing the length of audio messages," *International Journal of Human Computer Studies*, vol. 43, pp. 153-153, 1995.

[65] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "Experimentally derived guidelines for the creation of earcons," *Adjunct Proceedings of HCI*, vol. 95, pp. 155-159, 1995.

[66] E. Hoggan, A. Crossan, S. A. Brewster, and T. Kaaresoja, "Audio or tactile feedback: which modality when?," 2009.

[67] W. Barfield, C. Rosenberg, and G. Levasseur, "The use of icons, earcons, and commands in the design of an onlinehierarchical menu," *IEEE Transactions on Professional Communication*, vol. 34, pp. 101-108, 1991.

[68] S. Portigal, "Auralization of document structure," Citeseer, 1994.

[69] M. A. R. Kirby and D. Rigas, "Systems design: Ensuring the right users are involved."

[70] D. I. Rigas, M. A. R. Kirby, and D. O'Connel, "Audiotest: Utilising audio to communicate information in program debugging," 1999.

[71] D. Rigas and J. Alty, "The rising pitch metaphor: an empirical study," *International Journal of Human-Computer Studies*, vol. 62, pp. 1-20, 2005.

[72] S. Brewster and L. M. Brown, "Tactons: structured tactile messages for non-visual information display," 2004.

[73] D. I. Rigas and D. Memery, "Utilising audio-visual stimuli in interactive information systems: A two domain investigation on auditory metaphors," 2002.

[74] D. Rigas, H. Yu, and D. Memery, "Experiments using speech, non-speech sound and stereophony ascommunication metaphors in information systems," 2001.

[75] D. Rigas, D. Memery, and H. Yu, "Experiments in using structured musical sound, synthesised speech and environmental stimuli to communicate information: Is there a case for integration and synergy," *Proceedings of 2001 International Symbosium on Intelligent Multimedia Video and Speech Processing*, pp. 465–468, 2001.

[76] M. Forsberg, "Why is Speech Recognition Difficult?."

[77] R. A. Bolt, "“Put-that-there”: Voice and gesture at the graphics interface," 1980.

[78] L. R. Karl, M. Pettey, and B. Shneiderman, "Speech versus mouse commands for word processing: an empirical evaluation," 1998.

[79] V. Zue, J. Glass, D. Goddeau, D. Goodine, L. Hirschman, M. Phillips, J. Polifroni, and S. Seneff, "The MIT ATIS system: February 1992 progress report," 1992.

[80] A. K. Sinha, M. Shilman, and N. Shah, "MultiPoint: a case study of multimodal performance for building presentations," 2001.

[81] C. Benoit, J. C. Martin, C. Pelachaud, L. Schomaker, and B. Suhm, "Audio-visual and multimodal speech systems."

[82] D. H. Sonnenwald, B. Gopinath, G. O. Haberman, W. M. Keese Iii, J. S. Myers, and R. B. Bellcore, "InfoSound: An audio aid to program comprehension," 1990.

[83] K. F. Lee, H. W. Hon, and R. Reddy, "An overview of the SPHINX speech recognition system," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 38, pp. 35-45, 1990.

[84] D. Huggins-Daines, M. Kumar, A. Chan, A. W. Black, M. Ravishankar, and A. I. Rudnicky, "Pocketsphinx: A free, real-time continuous speech recognition system for hand-held devices," 2006.
[85] D. S. Weld, C. Anderson, P. Domingos, O. Etzioni, K. Gajos, T. Lau, and S. Wolfman, "Automatically personalizing user interfaces," 2003.
[86] S. Greenberg, *The computer user as toolsmith: The use, reuse, and organization of computer-based tools*: Cambridge Univ Pr, 1993.
[87] P. Langley, "User modeling in adaptive interfaces."
[88] A. Granic and V. Glavinic, "Adaptive systems and interaction: the design of personalized interaction in computer-based education," 2005.
[89] S. Perugini and N. Ramakrishnan, "Personalizing Web Sites with Mixed-Initiative Interaction," 2003.
[90] K. L. Wu, C. C. Aggarwal, and P. S. Yu, "Personalization with dynamic profiler," 2000.
[91] S. Braynov, "Personalization and customization technologies," *The Internet Encyclopedia*, vol. 3, 2003.
[92] M. Nadjarbashi-Noghani, J. Zhang, K. M. H. Sadat, and A. A. Ghorbani, "PENS: A personalized electronic news system," 2005.
[93] A. Pretschner and S. Gauch, "Personalization on the Web," 1999.
[94] J. McGrenere, "The design and evaluation of multiple interfaces: A solution for complex software," The University of Toronto, 2002.
[95] U. Manber, A. Patel, and J. Robison, "Experience with personalization of Yahoo!," 2000.
[96] A. Bunt, "Mixed-Initiative Support for Customizing Graphical User Interfaces," The University Of British Columbia, 2007.
[97] A. Bunt, G. Carenini, and C. Conati, "Adaptive Content Presentation for the Web," *The Adaptive Web: Methods and Strategies of Web Personalization, Lecture Notes in Computer Science*, vol. 4321.
[98] L. Console, I. Torre, I. Lombardi, S. Gioria, and V. Surano, "Personalized and adaptive services on board a car: An application for tourist information," *Journal of Intelligent Information Systems*, vol. 21, pp. 249-284, 2003.
[99] M. X. Zhou and V. Aggarwal, "An optimization-based approach to dynamic data content selection in intelligent multimedia interfaces," 2004.
[100] M. O'Donnell, C. Mellish, J. Oberlander, and A. Knott, "ILEX: An architecture for a dynamic hypertext generation system," *Natural Language Engineering*, vol. 7, pp. 225-250, 2001.
[101] J. R. Anderson, A. T. Corbett, K. R. Koedinger, and R. Pelletier, "Cognitive tutors: Lessons learned," *Journal of the Learning Sciences*, vol. 4, pp. 167-207, 1995.
[102] C. Conati and K. Vanlehn, "Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation," *International Journal of Artificial Intelligence in Education*, vol. 11, pp. 389-415, 2000.
[103] V. J. Shute and J. Psotka, "Intelligent tutoring systems: Past, present, and future," *Handbook of research on educational communications and technology*, pp. 570-600, 1996.
[104] K. R. Koedinger, J. R. Anderson, W. H. Hadley, and M. A. Mark, "Intelligent Tutoring Goes To School in the Big City," *International Journal of Artificial Intelligence in Education*, vol. 8, pp. 30-43, 1997.

[105] A. Bunt, C. Conati, and J. McGrenere, "Supporting interface customization using a mixed-initiative approach," *Proceedings of the 12th international conference on Intelligent user interfaces*, pp. 92-101, 2007.
[106] D. Billsus and M. J. Pazzani, "User modeling for adaptive news access," *User Modeling and User-Adapted Interaction*, vol. 10, pp. 147-180, 2000.
[107] D. Billsus and M. J. Pazzani, "A hybrid user model for news story classification," 1999.
[108] J. B. Schafer, J. A. Konstan, and J. Riedl, "E-Commerce Recommendation Applications."
[109] A. MacLean, K. Carter, L. Lövstrand, and T. Moran, "User-tailorable systems: pressing the issues with buttons," 1990.
[110] S. R. Page, T. J. Johnsgard, U. Albert, and C. D. Allen, "User customization of a word processor," 1996.
[111] M. Debevc, B. Meyer, D. Donlagic, and R. Svecko, "Design and evaluation of an adaptive icon toolbar," *User Modeling and User-Adapted Interaction*, vol. 6, pp. 1-21, 1996.
[112] S. R. Page, T. J. Johnsgard, U. Albert, and C. D. Allen, "User customization of a word processor," *Proceedings of the SIGCHI conference on Human factors in computing systems: common ground*, pp. 340-346, 1996.
[113] A. Sears and B. Shneiderman, "Split menus: Effectively using selection frequency to organize menus," 1998.
[114] K. Gajos and D. S. Weld, "Supple: automatically generating user interfaces," 2004.
[115] M. Gantt and B. A. Nardi, "Gardeners and gurus: patterns of cooperation among CAD users," 1992.
[116] B. E. John and D. E. Kieras, "Using GOMS for user interface design and evaluation: Which technique?," *ACM Transactions on Computer-Human Interaction*, vol. 3, pp. 287-319, 1996.
[117] G. Weber, "Adaptive learning systems in the World Wide Web," *COURSES AND LECTURES-INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES*, pp. 371-378, 1999.
[118] V. M. García-Barrios, F. Mödritscher, and C. Gütl, "Personalisation versus Adaptation? A User-centred Model Approach and its Application," 2005.
[119] L. Findlater, Moffatt, K., McGrenere, J., and Dawson, J. , "Ephemeral adaptation: The use of gradual onset to improve menu selection performance," *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI 2009)*, pp. 10, 2009.
[120] A. Holzinger, Kickmeier-Rust, M., & Albert, D., "Dynamic Media in Computer Science Education; Content Complexity and Learning Performance: Is Less More?," *Educational Technology & Society,*, vol. 11, pp. 279-290, 2008.
[121] S. Greenberg and I. H. Witten, "Adaptive personalized interfaces—A question of viability," *Behaviour & Information Technology*, vol. 4, pp. 31-45, 1985.
[122] C. de Beer, "Applying Artificial Intelligence Principles to Portal Customization-A Theoretical Approach."
[123] P. Brusilovsky, "Methods and techniques of adaptive hypermedia," *User modeling and user-adapted interaction*, vol. 6, pp. 87-129, 1996.
[124] T. Kühme, "A user-centered approach to adaptive interfaces," 1993.
[125] M. M. Tseng and F. T. Piller, *The customer centric enterprise: advances in mass customization and personalizaton*: Springer Verlag, 2003.

[126] B. Krulwich, "Lifestyle finder: Intelligent user profiling using large-scale demographic data," *AI magazine*, vol. 18, pp. 37, 1997.
[127] I. A. Al Mamunur Rashid, D. Cosley, S. K. Lam, S. M. McNee, J. A. Konstan, and J. Riedl, "Getting to know you: learning new user preferences in recommender systems," 2002.
[128] D. Greening, "What Marketers Really Want to Know: Tracking Users," *Web Techniques. July*, 1999.
[129] N. P. de Koch, "Software engineering for adaptive hypermedia systems."
[130] C. Stephanidis, C. Karagiannidis, and A. Koumpis, "Decision making in intelligent user interfaces."
[131] Y. Arens, E. Hovy, and S. Van Mulken, "Structure and rules in automated multimedia presentation planning," 1993.
[132] S. K. Feiner and K. R. McKeown, "Automating the generation of coordinated multimedia explanations," *Computer*, vol. 24, pp. 33-41, 1991.
[133] N. L. Green, G. Carenini, S. Kerpedjiev, J. Mattis, J. D. Moore, and S. F. Roth, "Autobrief: an experimental system for the automatic generation of briefings in integrated text and information graphics," *International Journal of Human-Computer Studies*, vol. 61, pp. 32-70, 2004.
[134] Y. Han and I. Zukerman, "Using Cooperative Agents to Plan Multimodal Presentations," *Lecture notes in computer science*, pp. 122-157, 1998.
[135] M. Maybury, "Planning multimedia explanations using communicative acts," *Readings in Intelligent User Interfaces*, pp. 107, 1998.
[136] R. F. Rodrigues, P. S. L. Rodrigues, B. Feijó, L. Velho, and L. F. G. Soares, "Cross-media and elastic time adaptive presentations: the integration of a talking head tool into a hypermedia formatter," *Lecture notes in computer science*, pp. 215-224, 2004.
[137] J. L. Herlocker, J. A. Konstan, and J. Riedl, "Explaining Collaborative Filtering Recommendations."
[138] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens: An open architecture for collaborative filtering of netnews."
[139] G. Linden and B. Smith, "Amazon. com recommendations: Item-to-item collaborative filtering," *IEEE Internet computing*, vol. 7, pp. 76-80, 2003.
[140] M. Eirinaki and M. Vazirgiannis, "Web mining for web personalization," *ACM Transactions on Internet Technology*, vol. 3, pp. 1-27, 2003.
[141] M. Balabanovic and Y. Shoham, "Content-Based, Collaborative Recommendation," *COMMUNICATIONS OF THE ACM*, vol. 40, pp. 67, 1997.
[142] P. Brusilovsky, "Adaptive educational systems on the world-wide-web: A review of available technologies," 1998.
[143] K. Z. Gajos, M. Czerwinski, D. S. Tan, and D. S. Weld, "Exploring the design space for adaptive graphical user interfaces," *Proceedings of the working conference on Advanced visual interfaces*, pp. 201-208, 2006.
[144] T. Tsandilas, "An Empirical Assessment of Adaptation Techniques," 2005.
[145] K. Z. Gajos, K. Everitt, D. S. Tan, M. Czerwinski, and D. S. Weld, "Predictability and accuracy in adaptive user interfaces," 2008.
[146] L. K. Silverman, "Identifying visual-spatial and auditory-sequential learners: A validation study," *Talent development V: Proceedings from the*, 2000.
[147] A. F. Smeaton and J. Callan, "Personalisation and recommender systems in digital libraries," *International Journal on digital libraries*, vol. 5, pp. 299-308, 2005.

[148] H. Hohl, H. D. Boecker, and R. Gunzenhaeuser, "Hypadapter: An adaptive hypertext system for exploratory learning and programming," *User Modeling and User-Adapted Interaction*, vol. 6, pp. 131-156, 1996.
[149] A. Micarelli and F. Sciarrone, "A case-based system for adaptive hypermedia navigation," *Lecture Notes in Computer Science*, pp. 266-279, 1996.
[150] A. Kobsa, D. Müller, and A. Nill, "KN-AHS: An adaptive hypertext client of the user modeling system BGP-MS," *Readings in intelligent user interfaces*, pp. 372, 1998.
[151] C. Boyle and A. O. Encarnacion, "MetaDoc: an adaptive hypertext reading system," *User Modeling and User-Adapted Interaction*, vol. 4, pp. 1-19, 1994.
[152] K. Höök, "Evaluating the utility and usability of an adaptive hypermedia system," *Knowledge-Based Systems*, vol. 10, pp. 311-319, 1998.
[153] F. Naumann, U. Leser, and J. C. Freytag, "Quality-driven integration of heterogeneous information systems," 1999.
[154] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," *Computer networks and ISDN systems*, vol. 30, pp. 107-117, 1998.
[155] R. Burke, "Hybrid recommender systems: Survey and experiments," *User Modeling and User-Adapted Interaction*, vol. 12, pp. 331-370, 2002.
[156] J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative Filtering Recommender Systems."
[157] M. J. Pazzani and D. Billsus, "Content-based recommendation systems," *Lecture Notes in Computer Science*, vol. 4321, pp. 325, 2007.
[158] R. Burke, "Knowledge-based recommender systems," *Encyclopedia of Library and Information Systems*, vol. 69, pp. 175-186, 2000.
[159] N. Good, J. B. Schafer, J. A. Konstan, A. Borchers, B. Sarwar, J. Herlocker, and J. Riedl, "Combining collaborative filtering with personal agents for better recommendations," 1999.
[160] P. Resnick and H. R. Varian, "Recommender systems," 1997.
[161] F. de Rosis, B. De Carolis, and S. Pizzutilo, "User tailored hypermedia explanations," 1993.
[162] T. Fox, G. Grunst, and K. J. Quast, "HyPLAN: A context-sensitive hypermedia help system," *Adaptive User Support. Ergonomic Design of Manually and Automatically Adaptable Software*, pp. 126-193, 1994.
[163] G. Fischer, "A critic for LISP," COLORADO UNIV AT BOULDER DEPT OF COMPUTER SCIENCE, 1987.
[164] L. M. Encarnação and S. L. Stoev, "An application-independent intelligent user support system exploiting action-sequence based user modelling," *COURSES AND LECTURES-INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES*, pp. 245-254, 1999.
[165] M. Gonschorek and C. Herzog, "Using hypertext for an adaptive helpsystem in an intelligent tutoring system," 1995.
[166] C. H. Chen, L. P. Khoo, and W. Yan, "PDCS—a product definition and customisation system for product concept development," *Expert Systems with Applications*, vol. 28, pp. 591-602, 2005.
[167] J. Wind and A. Rangaswamy, "Customerization: the next revolution in mass customization," *Journal of Interactive Marketing*, vol. 15, 2001.
[168] "The failure of customization: or why people don't buy jeans online," Knowledge@Wharton. University of Pennsylvania, March 2002.

[169] A. Bunt, C. Conati, and J. McGrenere, "What role can adaptive support play in an adaptable system?," *Proceedings of the 9th international conference on Intelligent user interfaces*, pp. 117-124, 2004.
[170] W. E. Mackay, "Triggers and barriers to customizing software," 1991.
[171] M. Pastore, "Customization Leads to E-Commerce Geographics," 1999.
[172] B. J. Pine and J. H. Gilmore, *The experience economy*: Harvard Business School Press Boston, MA, 1999.
[173] J. H. Gilmore and B. J. Pine, "The four faces of mass customization," *Harvard Business Review*, vol. 75, pp. 91-101, 1997.
[174] M. Sääksjärvi and T. Santonen, "Customization as business model for online newspapers," 2003.
[175] J. W. Palmer and L. B. Eriksen, "Digital newspapers explore marketing on the Internet," 1999.
[176] K. Haskell, "Do-It-Yourself Design: Just Point and Click," *The New York Times*, 2004.
[177] Y. J. Wind, "The challenge of “customerization” in financial services," 2001.
[178] J. Wind and V. Mahajan, "Digital marketing," 2001.
[179] W. E. Mackay, "Patterns of sharing customizable software," 1990.
[180] W. Stuerzlinger, O. Chapuis, D. Phillips, and N. Roussel, "User Interface Façades: Towards Fully Adaptable User Interfaces."
[181] B. Shneiderman, "Promoting universal usability with multi-layer interface design," 2003.
[182] L. Findlater and J. McGrenere, "Evaluating reduced-functionality interfaces according to feature findability and awareness," *LECTURE NOTES IN COMPUTER SCIENCE*, vol. 4662, pp. 592, 2007.
[183] G. Fischer, "Shared knowledge in cooperative problem-solving systems-integrating adaptive and adaptable components," *HUMAN FACTORS IN INFORMATION TECHNOLOGY*, vol. 10, pp. 49-49, 1993.
[184] E. Horvitz, "Principles of mixed-initiative user interfaces," 1999.
[185] C. G. Thomas and M. Krogsæter, "An adaptive environment for the user interface of Excel," *Proceedings of the 1st international conference on Intelligent user interfaces*, pp. 123-130, 1993.
[186] R. Oppermann, "Adaptive user support: ergonomic design of manually and automatically adaptable software," *Hillsdale, New Jersey, USA*, 1994.
[187] N. Lesh, C. Rich, and C. L. Sidner, "Using plan recognition in human-computer collaboration," *COURSES AND LECTURES-INTERNATIONAL CENTRE FOR MECHANICAL SCIENCES*, pp. 23-32, 1999.
[188] R. S. Amant and P. R. Cohen, "Interaction with a mixed-initiative system for exploratory data analysis," *Knowledge-Based Systems*, vol. 10, pp. 265-273, 1998.
[189] K. A. Papanikolaou, M. Grigoriadou, H. Kornilakis, and G. D. Magoulas, "Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE," *User Modeling and User-Adapted Interaction*, vol. 13, pp. 213-267, 2003.
[190] C. Price, A. Bunt, and G. McCalla, "L2tutor: A Mixed-Initiative Dialogue System for Improving Fluency," *Computer Assisted Language Learning*, vol. 12, pp. 83-112, 1999.
[191] S. A. Wolfman, T. Lau, P. Domingos, and D. S. Weld, "Mixed initiative interfaces for learning tasks: SMARTedit talks back," 2001.

[192] E. Horvitz, J. Breese, D. Heckerman, D. Hovel, and K. Rommelse, "The Lumiere project: Bayesian user modeling for inferring the goals and needs of software users," 1998.
[193] J. Koenemann and N. J. Belkin, "A case for interaction: a study of interactive information retrieval behavior and effectiveness," 1996.
[194] S. L. Jackson, J. Krajcik, and E. Soloway, "The design of guided learner-adaptable scaffolding in interactive learning environments," 1998.
[195] A. Cypher, "Eager: Programming repetitive tasks by example," 1991.
[196] B. Clark and J. Matthews, "Deciding layers: Adaptive composition of layers in a multi-layer user interface," 2005.
[197] A. Bunt, J. McGrenere, and C. Conati, "Understanding the Utility of Rationale in a Mixed-Initiative System for GUI Customization," *Proceedings of the International Conference on User Modeling (UM 2007)*, pp. 147–156, 2007.
[198] A. Jameson and E. Schwarzkopf, "Pros and Cons of Controllability: An Empirical Study," *Proc. Adaptive Hypermedia and Adaptive Web-Based Systems. Springer-Verlag, London*, pp. 193-202, 2002.
[199] R. Oppermann, "Adaptively supported adaptability," *International Journal of Human Computer Studies*, vol. 40, pp. 455-472, 1994.
[200] R. Coll and J. C. Wingertsman Iii, "The effect of screen complexity on user preference and performance," *International Journal of Human-Computer Interaction*, vol. 2, pp. 255-265, 1990.
[201] M. Shilman, D. S. Tan, and P. Simard, "CueTIP: a mixed-initiative interface for correcting handwriting errors," *Proceedings of the 19th annual ACM symposium on User interface software and technology*, pp. 323-332, 2006.
[202] I. Hsi and C. Potts, "Studying the Evolution and Enhancement of Software Features."
[203] F. D. Davis, R. P. Bagozzi, and P. R. Warshaw, "User acceptance of computer technology: a comparison of two theoretical models," *Management Science*, vol. 35, pp. 982-1003, 1989.
[204] X. Faulkner, *Usability engineering*: Houndmills, Basingstoke, Hampshire: Palgrave, 2000.
[205] E. Frekjmr and D. Copenhagen, "Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated?."

Appendices

Appendix A: Questionnaire of Experiment One

Questionnaire

We are investigating the use of different software interaction approaches and we would like to obtain your views about the use of such interactions. We compared between four conditions: Static, Customise, Personalise, and Mixed initiative (for more explanation see the terminology). Each was implemented in separate system.

Please follow the procedure bellow:

1. Answer Part 1 the Pre-Session questionnaire.
2. In Part 2 Read the scenario and the requirements of the tasks carefully.
3. Start the task.
4. On completion of the task or when requested to stop, answer the questions provided from the point of view of customer.
5. Do the same procedure from step 2, 3, and 4 for the tasks in Part 3, and Part 4.
6. On completion of the tasks, answer the questions provided in Part 5.
7. Do the same procedure from step 2, 3, and 4 for the tasks in Part 6.
8. On completion of the tasks, answer the questions provided.
9. For customisation and Mixed-initiative procedure provided at the end of this document.

Terminology

- **Static**: The interface and content does not change over time.
- **Customisation**: The interface and content change over time by user.
- **Personalisation**: The interface and content change over time by system.
- **Mixed-initiative**: The interface and content change over time by user and system.

It would be grateful if you could fill in the following questionnaire sincerely and provide your views. Thank you for your participation.

Khalid Al-Omar

Part 1	Pre-Session Questionnaire

1. What is your gender?
   - ☐ Male ☐ Female

2. In what age group are you?
   - ☐ 17 and under ☐ 18-24 ☐ 25-34
   - ☐ 35-44 ☐ 45-54 ☐ 55 +

3. What is your education level?
   - ☐ High school ☐ College degree ☐ Graduate
   - ☐ Post-graduate ☐ Other [ ]

4. How often do you use a computer (average)?
   - ☐ Never ☐ Rarly (1-2 times in 3 months)
   - ☐ Occasionally (1-2 times a month) ☐ 1-2 times a week
   - ☐ 3-5 times a week ☐ More than 5 times a week

5. How many hours you use the internet (average) per week?
   - ☐ Less than 1 hour ☐ 1-5 a hours ☐ 6-10 hours
   - ☐ More than 10 hours

6. How often do you shop online?
   - ☐ Never ☐ Occasionally (1-2 times a month)
   - ☐ 1-2 times a week ☐ 3-5 times a week ☐ More than 5 times a week

7. Which type of programming language you used?
   - ☐ Java ☐ C#
   - ☐ VB.Net ☐ PHP ☐ Ada
   - ☐ HTML ☐ None

8. Do you ever change software setting?
   - ☐ No, never
   - ☐ Yes, every time I use a new software
   - ☐ Yes, when I need too ☐ Yes, when I get some errors

9. In the internet, do you ever use any customisable web page such as “My Yahoo” or “Google Personalised home page”?
   - ☐ No, never ☐ Yes, once ☐ Sometimes (1-2 times) ☐ Everytime

Part 2	Tasks 1,2, and 3

Scenario: Assume that you are vegetarian and you are allergic to eggs, milk, and alcholic drinks. Imagine that you want to purchase some different type of fruit and vegetables and you wish not see the items that you allergic to. To accomplish task 1, 2, and 3, you need to Register with the system first and you should consider this summary:

Go for	Avoid
Fruit and Vegetables.	Meat Fish Bacon Eggs Milk Alcholic drinks

Task 1: Go to " Frozen Food" and purchase one vegetarian item.

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..

Task 2: Go to " Fresh Food" and purchase one vegetarian item.

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..

Task 3: Go to "Ready Meal" and purchase one vegetable ready meal item and then go to "Fresh Food" and purchased one vegetarian item that you purchased before.

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Part 2	Questionnaire

2.1 Below is a list of questions. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Extremely easy	Very easy	Easy	Not very easy	Not at all easy
1. How easy was using the website?	1	2	3	4	5
2. How easy was performing the tasks?	1	2	3	4	5
3. How easy you felt searching for items?	1	2	3	4	5
	Extremely comfortable	Very uncomfort	comfort	Not very comfort	Not at comfort
4. How comfortable you felt by using the keyboard?	1	2	3	4	5
5. How confident you felt using the website?	1	2	3	4	5

2.2 Below is a list of statements. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Strongly Disagree	Disagree	Mildly Disagree	Mildly Agree	Agree	Strongly Agree
6. The contents of the website are very suitable.	1	2	3	4	5	6
7. It was easy to make the website do exactly what I want.	1	2	3	4	5	6
8. I think the website is confusing.	1	2	3	4	5	6
9. There have been times in using the website when I have felt nervous.	1	2	3	4	5	6
10. I will never learn to use all features that offered in this website.	1	2	3	4	5	6
11. I feel the website help me to do my shopping.	1	2	3	4	5	6
12. Over all I like this website.	1	2	3	4	5	6

2.3 Rate level of effect of the following reasons on your selecetion of purchased.

	the lowest									the highest
13. I like these items:	1	2	3	4	5	6	7	8	9	10
14. Items position on the list convincing:	1	2	3	4	5	6	7	8	9	10

Part 3	Tasks 4, 5, and 6

Scenario: Assume that you are vegetarian and you have a dinner party. You invited three of your friends and they are not vegeterian. So you been asked to purchase items that you are allergic to. To accomplish task 4, 5, and 6, you need to Register with the system first and you should consider this summary:

Go for	Avoid
- Fruit and Vegetables. - Meat - Fish - Bacon - Eggs - Milk - Alcholic drinks - High Energy	---

Task 4: Purchase the following items:

Number	Category	Item name	Notes	By
1	Fresh Food	Smoked Salmon Organic	cheapest	Keyboard
2	Ready Meal	Chicken with wine sauce	cheapest	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Task 5: Purchase the following items:

Number	Category	Item name	Notes	By
1	Drinks	Milk	cheapest	cliks
2	Drinks	Soft Drink	The highest energy drink	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Task 6: Purchase the following items:

Number	Category	Item name	Notes	By
1	Ice cream	Rice Puddings	Change quntity 2	Keyboard
2	Ice cream	Rice Puddings	with Strawberry	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Part 3	**Questionnaire**

3.1 Below is a list of questions. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Extremely easy	Very easy	Easy	Not very easy	Not at all easy
1. How easy was using the website?	1	2	3	4	5
2. How easy was performing the tasks?	1	2	3	4	5
3. How easy you felt searching for items?	1	2	3	4	5

	Extremely comfortable	Very uncomfort	comfort	Not very comfort	Not at comfort
4. How comfortable you felt by using the keyboard?	1	2	3	4	5
5. How confident you felt using the website?	1	2	3	4	5

3.2 Below is a list of statements. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Strongly Disagree	Disagree	Mildly Disagree	Mildly Agree	Agree	Strongly Agree
6. The contents of the website are very suitable.	1	2	3	4	5	6
7. It was easy to make the website do exactly what I want.	1	2	3	4	5	6
8. I think the website is confusing.	1	2	3	4	5	6
9. There have been times in using the website when I have felt nervous.	1	2	3	4	5	6
10. I will never learn to use all features that offered in this website.	1	2	3	4	5	6
11. I feel the website help me to do my shopping.	1	2	3	4	5	6
12. Over all I like this website.	1	2	3	4	5	6

3.3 Rate level of effect of the following reasons on your selecetion of purchased.

	the lowest									the highest
13. I like these items:	1	2	3	4	5	6	7	8	9	10
14. Items position on the list convincing:	1	2	3	4	5	6	7	8	9	10

Part 4	Tasks 7, 8, and 9

Scenario: You have a high blode pressure and a high Cholestrol levels. Your doctor recommended to eat some protien-rich foods such as meat, fish, and eggs and to avoid food contains high Fats, Carbohydrate, Sugar, and Salt. To accomplish tasks 7, 8, and 9, you need to Register with the system first and you should consider this summary:

Go for	Avoid
- Fruit and Vegetables. - Eggs and Milk - Meat - High Protien - Organic food	- High Fats. - High Carbohydrate. - High Sugar. - High Salt.

Task 7: Purchase the following items:

Item Number	Item name	Nutrition	Notes	By
1	Smoked Salmon	High Protien more than 2.9 g	cheapest	Keyboard
2	Pork	High Protien more than 2.9 g	cheapest	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Task 8: Purchase the following items:

Number	Item name	Nutrition	Notes	By
1	Cheese	Low Fat	cheapest	Keyboard
2	Soft Drinks	Energy = 122 Kcal Protien < 2.9 g	-	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Task 9: Purchase the following items:

Number	Item name	Nutrition	Notes	By
1	Cheese	Energy >=122 Kcal Protien = 2.9 g	cheapest	Keyboard
2	Chicken	Protien = 2.9 g Carbohydrate = 9.6 g Salt < Fats <	change quntity to 2	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Part 4	**Questionnaire**

4.1 Below is a list of questions. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Extremely easy	Very easy	Easy	Not very easy	Not at all easy
1. How easy was using the website?	1	2	3	4	5
2. How easy was performing the tasks?	1	2	3	4	5
3. How easy you felt searching for items?	1	2	3	4	5

	Extremely comfortable	Very uncomfort	comfort	Not very comfort	Not at comfort
4. How comfortable you felt by using the keyboard?	1	2	3	4	5
5. How confident you felt using the website?	1	2	3	4	5

4.2 Below is a list of statements. We would like you to tell us what your attitude is to each by circling one of the numbers.

	Strongly Disagree	Disagree	Mildly Disagree	Mildly Agree	Agree	Strongly Agree
6. The contents of the website are very suitable.	1	2	3	4	5	6
7. It was easy to make the website do exactly what I want.	1	2	3	4	5	6
8. I think the website is confusing.	1	2	3	4	5	6
9. There have been times in using the website when I have felt nervous.	1	2	3	4	5	6
10. I will never learn to use all features that offered in this website.	1	2	3	4	5	6
11. I feel the website help me to do my shopping.	1	2	3	4	5	6
12. Over all I like this website.	1	2	3	4	5	6

4.3 Rate level of effect of the following reasons on your selecetion of purchased.

	the lowest									the highest
13. I like these items:	1	2	3	4	5	6	7	8	9	10
14. Items position on the list convincing:	1	2	3	4	5	6	7	8	9	10

Part 5	Task 10, 11 and 12

Go for	Avoid
Fruit and Vegetables.	Meat Fish Bacon Eggs Milk Alcholic drinks

Task 10: Go to " Frozen Food" and purchase one vegetarian item.

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..

Task 11: Purchase the following items:

Number	Category	Item name	Notes	By
1	Fresh Food	Smoked Salmon Organic	cheapest	Keyboard
2	Ready Meal	Chicken with wine sauce	cheapest	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Task 12: Purchase the following items:

Item Number	Item name	Nutrition	Notes	By
1	Smoked Salmon	High Protien more than 2.9 g	cheapest	Keyboard
2	Pork	High Protien more than 2.9 g	cheapest	Keyboard

Answer:

No	Item name	Item Id
1.	………………………………………..	…………..
2.	………………………………………..	…………..

Part 5	Task 13, and 14

Task 13: Choose the approache that you like most and purchase as many items as you like:

Answer:

Number of purchased item:……………………

1) Rate level of effect of the following reasons on your selecetion of purchased.

	the lowest									the highest
1. I like these items:	1	2	3	4	5	6	7	8	9	10
2. Items position on the list convincing:	1	2	3	4	5	6	7	8	9	10
3. The website is convincing:	1	2	3	4	5	6	7	8	9	10

Task 14: Choose the approache that you dislike most and purchase as many items as you like.

Answer:

Number of purchased item:……………………

1) Rate level of effect of the following reasons on your selecetion of purchased.

	the lowest									the highest
1. I like these items:	1	2	3	4	5	6	7	8	9	10
2. Items position on the list convincing:	1	2	3	4	5	6	7	8	9	10
3. The website is convincing:	1	2	3	4	5	6	7	8	9	10

Part 6	Post-Experimental Views

Please answer the following questions.

1. Which approach did you prefer overall and why? Please put a 1 beside that approach and a 2 beside the next preferred approach.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

2. With which approach did you find the easiest to use? Please put a 1 beside that approach and a 2 beside the second easiest to use.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

3. With which approach did you feel you were the most efficient (quick)? Please put a 1 beside that approach and a 2 beside the next most efficient.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

4. With which approach did you feel the most frustrated? Please put a 1 beside that approach and a 2 beside the next most confident.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

5. With which approach did you prefer the most Interfaces? Please put a 1 beside that approach and a 2 beside the next most preferable interface.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

6. With which approach did you think that it helps you with your shopping? Please put a 1 beside that approach and a 2 beside the next most helpful approach.

   __ Static __ Personalisation __ Customisation __Mixed-initiative

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

   Did you find any difference between the four approaches and why?

   ☐ Yes ☐ No

   Why? ……………………………………………………………………………
   ……………………………………………………………………………

Procedure for Customisation and Mixed-initiative

1. Task 4, 5, and 6.

To customise Task 4, 5, and 6 do the following:

1. Go to “Fresh Food”.
2. Select “Sea Food”.
3. Make all “Smoked Salmon Organic” at the top of the list.
4. Go to “Ready Meal”.
5. Select “Traditional”.
6. Make all “Chicken with wine” at the top of the list.
7. Go to “Drinks”.
8. Select “Soft Drinks”.
9. Make all “Soft Drinks” at the top of the list.
10. Go to “Drinks”.
11. Select “Milk”.
12. Make all “Milk” at the top of the list.
13. Sort all items by price from low to high.
14. Go to “Frozen Food”.
15. Select “Ice cream”.
16. Make all “Rice Puddings” at the top of the list.
17. Finally customise the keyboard in the way is applicable to you.

2. Task 7, 8, and 9.

To customise Task 7, 8, and 9 do the following:

1. Go to “Fresh Food”.
2. Select “Sea Food”.
3. Sort all “Smoked Salmon” at the top of the list.
4. Click “Add new content”.
5. Click “Content List”.
6. Choose “Cheese”.
7. Add to “WebPartZone1”.
8. Click “Browse Mode”.
9. Finally customise the keyboard in the way is applicable to you.

3. Task 10, 11, and 12.

To customise Task 10, 11, and 12 do the following:

1. Go to “Fresh Food”.
2. Select “Sea Food”.
3. Make all “Smoked Salmon Organic” at the top of the list.
4. Go to “Ready Meal”.
5. Select “Traditional”.
6. Make all “Chicken with wine” at the top of the list.
7. Go to “Fresh Food”.
8. Select “Sea Food”.
9. Make all “Smoked Salmon” at the top of the list.
10. Finally customise the keyboard in the way is applicable to you.

Appendix B: Row Data from Experiment One

B.1 Task accomplishment time (related to section 3.8.1.1)

B.1.1 Task accomplishment time for Adaptive condition

Adaptive											
	Easy			Medium			Complex				
User no	TE1	TE2	TE3	TM1	TM2	TM3	C1	C2	C3	Average	Sum
User 1	20	12	85	55	65	55	46	323	78	82.11	739
User 2	27	16	17	43	73	20	60	272	205	81.44	733
User 3	29	13	15	23	25	25	43	235	182	65.55	590
User 4	165	22	16	32	28	24	160	172	222	93.44	841
User 5	23	170	45	31	25	31	124	170	606	136.11	1225
User 6	124	10	242	21	17	18	205	158	242	115.22	1037
User 7	11	10	19	76	98	85	106	136	238	86.55	779
User 8	8	17	66	25	97	180	91	110	356	105.55	950
User 9	37	13	36	63	25	38	50	323	322	100.77	907
User 10	20	16	10	41	50	18	163	272	182	85.77	772
User 11	124	13	32	28	73	52	134	158	222	92.88	836
User 12	8	56	22	55	43	24	45	172	606	114.55	1031
User 13	11	31	85	43	28	155	137	170	210	96.66	870
User 14	165	56	122	34	54	33	104	158	238	107.11	964
User 15	23	22	19	43	59	23	102	136	182	67.66	609

B.1.2 Task accomplishment time for Adaptable condition

Adaptable											
	Easy			Medium			Complex				
User no	TE1	TE2	TE3	TM1	TM2	TM3	C1	C2	C3	Average	Sum
User 1	10	24	34	63	49	182	266	247	209	120.44	1084
User 2	34	29	31	121	35	37	38	73	141	59.88	539
User 3	22	15	33	27	40	87	100	70	137	59	531
User 4	31	26	69	220	200	155	84	205	352	149.11	1342
User 5	32	21	117	180	55	117	93	186	596	155.22	1397
User 6	10	24	34	63	49	182	85	209	254	101.11	910
User 7	25	22	53	81	55	97	159	259	129	97.77	880
User 8	22	21	29	221	43	30	171	102	394	114.77	1033
User 9	10	16	62	92	57	58	172	249	203	102.11	919
User 10	22	17	56	68	39	41	179	249	243	101.55	914
User 11	31	21	32	105	41	105	134	259	223	105.67	951
User 12	31	22	73	70	62	127	170	186	411	128	1152
User 13	25	18	57	122	89	98	97	249	162	101.88	917
User 14	22	23	31	154	48	56	121	209	269	103.66	933
User 15	21	15	54	121	71	102	143	249	208	109.33	984

B.1.3 Task accomplishment time for Static condition

Static											
	Easy			Medium			Complex				
User no	TE1	TE2	TE3	TM1	TM2	TM3	C1	C2	C3	Average	Sum
User 1	54	40	80	186	69	57	108	287	323	133.77	1204
User 2	34	66	76	65	34	34	132	466	318	136.11	1225
User 3	38	33	116	169	105	140	156	264	600	180.11	1621
User 4	26	104	54	127	135	114	138	214	738	183.33	1650
User 5	15	50	55	104	117	122	420	161	1440	276	2484
User 6	47	36	73	84	144	143	192	348	757	202.66	1824
User 7	63	63	78	161	182	211	405	353	1260	308.44	2776
User 8	38	67	112	65	181	267	165	416	462	197	1773
User 9	20	27	51	142	113	163	214	277	384	154.55	1391
User 10	26	100	99	125	27	38	45	50	265	86.11	775
User 11	49	105	149	141	57	87	163	302	1093	238.44	2146
User 12	56	40	37	246	40	35	138	196	303	121.22	1091
User 13	27	30	76	85	105	143	112	453	669	188.88	1700
User 14	30	38	49	42	196	161	137	313	272	137.56	1238
User 15	50	20	73	97	139	131	123	343	360	154.87	1239

B.1.4 Task accomplishment time for Mixed-initiative condition

Mixed-Initiative											
	Easy			Medium			Complex				
User no	TE1	TE2	TE3	TM1	TM2	TM3	C1	C2	C3	Average	Sum
User 1	20	27	20	56	62	17	68	25	112	45.22	407
User 2	20	20	19	60	46	80	96	236	110	76.33	687
User 3	18	20	15	64	39	21	240	151	195	84.77	763
User 4	33	15	14	85	38	22	212	158	100	75.22	677
User 5	23	18	24	85	51	51	65	214	195	80.66	726
User 6	57	40	27	28	53	23	95	206	196	80.55	725
User 7	20	27	15	66	42	16	162	230	253	92.33	831
User 8	30	15	20	14	20	29	102	129	274	70.33	633
User 9	19	22	18	55	31	39	184	168	232	85.33	768
User 10	37	21	17	40	47	32	107	216	241	84.22	758
User 11	25	24	21	60	37	22	82	193	187	72.33	651
User 12	27	22	19	66	41	50	68	152	100	60.55	545
User 13	20	27	18	56	50	19	202	32	114	59.77	538
User 14	30	18	27	85	49	47	95	157	212	80	720
User 15	30	20	15	35	38	19	180	192	116	71.66	645

B.1.5 Task accomplishment time for training tasks of Adaptive condition

	Adaptive All Training Tasks				
User no	Easy L	Medium L	Complex L	Average	Sum
User 1	10	23	62	32	95
User 2	12	25	172	70	209
User 3	22	18	160	67	200
User 4	18	14	164	65	196
User 5	15	18	223	85	256
User 6	186	20	66	91	272
User 7	42	72	78	64	192
User 8	3	23	25	17	51
User 9	27	13	98	46	138
User 10	80	26	56	54	162
User 11	12	36	151	66	199
User 12	24	22	139	62	185
User 13	67	19	130	72	216
User 14	45	18	124	62	187
User 15	52	14	138	68	204

B.1.6 Task accomplishment time for training tasks of Adaptable condition

Adaptable All Training Tasks					
User no	Easy L	Medium L	Complex L	Average	Sum
User 1	26	55	51	44	132
User 2	27	50	51	43	128
User 3	18	26	41	28	85
User 4	32	52	52	45	136
User 5	26	60	44	43	130
User 6	26	55	51	44	132
User 7	25	24	107	52	156
User 8	62	56	56	58	174
User 9	22	16	21	20	59
User 10	19	38	73	43	130
User 11	32	45	56	44	133
User 12	28	55	47	43	130
User 13	23	33	62	39	118
User 14	18	28	51	32	97
User 15	26	52	44	41	122

B.1.7 Task accomplishment time for training tasks of Mixed-initiative condition

	Mixed All Training Tasks				
User no	Easy L	Medium L	Complex L	Average	Sum
User 1	30	24	23	26	77
User 2	13	23	133	56	169
User 3	17	22	60	33	99
User 4	12	100	40	51	152
User 5	19	43	88	50	150
User 6	30	14	54	33	98
User 7	17	12	61	30	90
User 8	22	24	73	40	119
User 9	30	24	23	26	77
User 10	13	23	75	37	111
User 11	17	22	60	33	99
User 12	12	100	40	51	152
User 13	18	43	88	50	149
User 14	29	14	54	32	97
User 15	15	12	61	29	88

B.1.8 Task accomplishment time for training tasks of Static condition

	Static All Training Tasks				
User no	Easy L	Medium L	Complex L	Average	Sum
User 1	19	55	48	41	122
User 2	65	131	76	91	272
User 3	14	67	56	46	137
User 4	32	58	61	50	151
User 5	51	66	100	72	217
User 6	88	42	82	71	212
User 7	40	105	103	83	248
User 8	45	75	37	52	157
User 9	26	48	31	35	105
User 10	25	93	123	80	241
User 11	66	85	175	109	326
User 12	56	45	95	65	196
User 13	20	86	151	86	257
User 14	43	56	156	85	255
User 15	22	56	100	59	178

B.1.9 Task accomplishment time for keyboard of Adaptive condition

Adaptive								
	Medium Tasks				Complex Tasks			
User no	T1	T2	T3	Average	T1	T2	T3	Average
User1	24	57	43	41	61	59	19	46
User2	35	44	38	39	28	42	43	38
User3	27	41	21	29	52	30	26	36
User4	17	18	9	14	22	26	35	28
User5	20	32	12	21	20	26	28	25
User6	15	19	22	18	95	26	28	50
User7	50	29	16	31	67	39	63	56

User8	13	91	58	54	61	21	78	53
User9	19	29	21	23	28	59	19	35
User10	25	35	60	40	52	42	43	46
User11	19	57	12	29	22	30	26	26
User12	27	44	22	31	20	26	35	27
User13	35	41	16	30	95	26	28	50
User14	17	18	58	31	67	26	28	40
User15	25	32	21	26	28	39	63	43
Average	24.53	39.13	28.60		47.87	34.47	37.47	
Median	24	35	21		52	30	28	
Mode	35	57	21		28	26	28	
Min	13	18	9		20	21	19	
Max	50	91	60		95	59	78	
SD	10.3	20.5	19.4		26.2	12.0	17.6	

B.1.10 Task accomplishment time for keyboard of Adaptable condition

Adaptable								
	Medium Tasks				Complex Tasks			
User no	T1	T2	T3	Average	T1	T2	T3	Average
User1	16	10	10	12	14	16	56	29
User2	12	69	30	37	30	11	29	23
User3	14	55	15	28	15	13	15	14
User4	20	26	39	28	17	17	38	24
User5	30	24	62	39	9	60	65	45
User6	16	69	10	32	14	16	56	29
User7	21	15	66	34	49	26	32	36
User8	42	21	18	27	44	19	35	33
User9	23	15	6	15	18	17	18	18
User10	28	32	13	24	22	22	26	23
User11	42	69	10	40	49	60	65	58
User12	23	15	66	35	44	16	56	39
User13	28	21	18	22	18	26	32	25
User14	16	15	6	12	22	19	35	25
User15	21	32	13	22	14	60	18	31
Average	23.47	32.53	25.47		25.27	26.53	38.40	
Median	21	24	15		18	19	35	
Mode	16	15	10		14	16	56	
Min	23.47	32.53	25.47		9	11	15	
Max	21	24	15		49	60	65	
SD	8.7	19.4	24.2		14.1	17.8	17.1	

B.1.11 Task accomplishment time for keyboard of Mixed-initiative condition

Mixed-initiative								
	Medium Tasks				Complex Tasks			
User no	T1	T2	T3	Average	T1	T2	T3	Average
User1	23	35	9	22	27	23	28	26
User2	26	69	13	36	25	27	48	33
User3	25	47	10	27	26	19	41	29
User4	24	60	18	34	27	18	32	26
User5	78	61	51	63	26	19	41	29
User6	20	41	11	24	13	47	43	34
User7	33	40	33	35	47	40	60	49
User8	12	22	8	14	14	23	40	26
User9	23	35	9	22	27	23	28	26

User10	26	69	13	36	25	27	48	33
User11	25	47	10	27	26	19	41	29
User12	24	60	18	34	27	18	32	26
User13	78	61	51	63	26	19	41	29
User14	20	41	11	24	13	47	43	34
User15	33	40	33	35	47	40	60	49
Average	31.33	48.53	19.87		26.40	27.27	41.73	
Median	25	47	13		26	23	41	
Mode	23	35	9		27	19	41	
Min	12	22	8		13	18	28	
Max	78	69	51		47	47	60	
SD	21.8	13.9	15.9		9.9	10.7	9.7	

B.1.12 Task accomplishment time for keyboard of Static condition

Static								
	Medium Tasks				Complex Tasks			
User no	T1	T2	T3	Average	T1	T2	T3	Average
User1	28	25	25	26	18	32	16	22
User2	38	23	48	36	32	32	43	36
User3	28	36	30	32	28	22	45	32
User4	52	45	19	39	28	22	20	23
User5	11	23	16	17	17	62	50	43
User6	15	158	18	64	23	23	35	27
User7	19	31	25	25	20	23	21	21
User8	23	39	25	29	23	29	68	40
User9	44	34	70	49	33	25	37	32
User10	23	28	21	24	21	27	34	27
User11	19	27	24	23	27	30	46	34
User12	25	34	26	29	55	52	33	47
User13	23	39	93	51	30	30	50	37
User14	19	64	26	36	55	52	35	47
User15	19	34	93	49	30	30	21	27
Average	25.75	42.70	37.29		29.36	32.73	36.87	
Median	23	34.34	25		27.89	29.61	35	
Mode	19	34.34	25		27.89	22	50	
Min	11	22.7	16		17	22	15.68	
Max	52	158	93		55.21	62	68	
SD	11.8	36.7	29.2		11.6	12.4	14.0	

B.2 Number of clicks and visited pages (related to section 3.8.1.2)

B.2.1 Number of clicks and visited pages for Adaptive condition

	Easy									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	2	1	5	8	2.6	3	2	3	8	2.6
User 2	2	1	1	4	1.3	2	2	2	6	2
User 3	2	1	1	4	1.3	2	2	2	6	2
User 4	9	2	1	12	4	2	2	2	6	2
User 5	3	18	4	25	8.33	2	11	2	15	5
User 6	15	2	26	43	14.33	8	2	7	17	5.6
User 7	1	2	1	4	1.33	2	2	2	6	2
User 8	3	2	3	8	2.6	2	2	2	6	2
User 9	4	1	4	9	3	2	2	2	6	2
User 10	1	2	1	4	1.33	2	2	2	6	2
User 11	15	10	1	26	8.6	8	10	2	20	6.6
User 12	3	2	4	9	3	2	2	2	6	2
User 13	1	2	1	4	1.33	2	2	2	6	2
User 14	9	2	15	26	8.6	2	2	7	11	3.6
User 15	3	1	3	7	2.33	2	2	2	6	2
Average	4.87	3.27	4.73	12.87	4.29	2.87	3.13	2.73	8.73	2.91
Median	3	2	3	8	2.66	2	2	2	6	2
Mode	3	2	1	4	1.33	2	2	2	6	2
Min	1	1	1	4	1.33	2	2	2	6	2
Max	15	18	26	43	14.33	8	11	7	20	6.66

	Medium									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	4	6	3	13	4	3	3	3	9	3
User 2	3	2	1	6	2	2	2	2	6	2
User 3	1	1	1	3	1	2	2	2	6	2
User 4	3	2	1	6	2	2	2	2	6	2
User 5	4	3	3	10	3	2	2	2	6	2
User 6	2	1	1	4	1	2	2	2	6	2
User 7	5	2	8	15	5	2	2	3	7	2
User 8	1	2	17	20	7	2	2	7	11	4
User 9	7	1	1	9	3	2	2	2	6	2
User 10	4	6	3	13	4	3	3	3	9	3
User 11	3	2	8	13	4	2	2	3	7	2
User 12	1	1	1	3	1	2	2	2	6	2
User 13	4	3	17	24	8	2	2	7	11	4
User 14	2	1	3	6	2	2	2	2	6	2
User 15	5	2	1	8	3	2	2	2	6	2
Average	3.27	2.33	4.60	10.20	3.40	2.13	2.13	2.93	7.20	2.40
Median	3	2	3	9	3	2	2	2	6	2
Mode	4	2	1	13	4.33	2	2	2	6	2
Min	1	1	1	3	1	2	2	2	6	2
Max	7	6	17	24	8	3	3	7	11	3.67

	Complex									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	1	23	10	34	11	2	11	4	17	6
User 2	3	2	12	17	6	3	27	5	35	12
User 3	1	16	2	19	6	2	6	3	11	4
User 4	14	15	17	46	15	11	10	5	26	9
User 5	15	18	60	93	31	8	11	42	61	20
User 6	15	9	26	50	17	8	5	7	20	7
User 7	9	12	14	35	12	5	4	5	14	5
User 8	1	2	25	28	9	2	2	10	14	5
User 9	1	16	6	23	8	2	6	4	12	4
User 10	14	15	17	46	15	11	10	5	26	9
User 11	15	15	60	90	30	8	11	42	61	20
User 12	15	9	26	50	17	8	5	7	20	7
User 13	9	12	14	35	12	5	4	5	14	5
User 14	1	2	17	20	7	2	2	5	9	3
User 15	1	12	6	19	6	2	4	4	10	3
Average	7.67	11.87	20.80	40.33	13.44	5.27	7.87	10.20	23.33	7.78
Median	9	12	17	35	11	5	6	5	17	5
Mode	1	2	17	19	6	2	11	5	14	4
Min	1	2	2	17	5	2	2	3	9	3
Max	15	23	60	93	31	11	27	42	61	20

B.2.2 Number of clicks and visited pages for Adaptable condition

	Easy									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	3	4	4	11	4	2	2	2	6	2
User 2	3	2	3	8	3	4	2	4	10	3
User 3	6	3	4	13	4	2	2	2	6	2
User 4	4	3	6	13	4	4	2	3	9	3
User 5	4	4	12	20	7	2	1	3	6	2
User 6	3	4	4	11	4	2	2	2	6	2
User 7	5	4	3	12	4	2	2	2	6	2
User 8	6	4	3	13	4	2	2	4	8	3
User 9	4	4	5	13	4	2	2	4	8	3
User 10	4	4	7	15	5	3	3	4	10	3
User 11	4	4	3	11	4	4	3	4	11	4
User 12	4	4	6	14	5	4	2	3	9	3
User 13	3	3	3	9	3	2	2	2	6	2
User 14	6	4	3	13	4	2	2	4	8	3
User 15	3	3	3	9	3	2	2	2	6	2
Average	4.13	3.60	4.60	12.33	4.11	2.60	2.07	3.00	7.67	2.56
Median	4	4	4	13	4.3	2	2	3	8	2.6
Mode	4	4	3	13	4.3	2	2	2	6	2
Min	3	2	3	8	2.6	2	1	2	6	2
Max	6	4	12	20	6.6	4	3	4	11	3.6

	Medium									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	3	4	21	28	9	2	2	7	11	4
User 2	14	3	4	21	7	8	2	4	14	5
User 3	4	3	16	23	8	2	2	9	13	4
User 4	23	20	16	59	20	8	2	12	22	7
User 5	14	4	12	30	10	6	2	9	17	6
User 6	3	4	21	28	9	2	2	7	11	4
User 7	4	4	9	17	6	2	2	9	13	4
User 8	19	2	2	23	8	7	2	1	10	3
User 9	7	3	4	14	5	4	1	3	8	3
User 10	6	3	6	15	5	4	1	5	10	3
User 11	3	3	9	15	5	2	1	9	12	4
User 12	14	4	12	30	10	8	2	9	19	6
User 13	4	4	9	17	6	2	2	9	13	4
User 14	23	4	4	31	10	8	2	3	13	4
User 15	3	4	9	16	5	2	2	9	13	4
Average	9.60	4.60	10.27	24.47	8.16	4.47	1.80	7.00	13.27	4.42
Median	6.00	4.00	9.00	23.00	7.67	4.00	2.00	9.00	13.00	4.33
Mode	3.00	4.00	9.00	28.00	9.33	2.00	2.00	9.00	13.00	4.33
Min	3.00	2.00	2.00	14.00	4.67	2.00	1.00	1.00	8.00	2.67
Max	23.00	20.00	21.00	59.00	19.67	8.00	2.00	12.00	22.00	7.33

	Complex									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	7	27	14	48	16	4	7	10	21	7
User 2	3	4	26	33	11	2	4	25	31	10
User 3	11	8	14	33	11	10	6	6	22	7
User 4	9	32	30	71	24	8	29	17	54	18
User 5	11	32	73	116	39	10	24	31	65	22
User 6	9	12	14	23	12	8	11	10	18	9
User 7	16	26	6	48	16	13	22	6	41	14
User 8	14	6	19	39	13	7	5	16	28	9
User 9	31	29	22	82	27	20	23	16	59	20
User 10	15	21	18	54	18	10	19	14	43	14
User 11	14	32	14	60	20	7	24	10	41	14
User 12	12	29	53	94	31	10	21	31	62	21
User 13	11	26	6	43	14	9	22	6	37	12
User 14	11	25	22	58	19	9	21	16	46	15
User 15	16	4	18	38	13	13	4	14	31	10
Average	12.7	21.5	23.3	56.0	18.9	9.3	16.5	15.2	39.9	13.5
Median	11.0	26.0	18.0	48.0	16.0	9.0	21.0	14.0	41.0	13.7
Mode	11.0	32.0	14.0	48.0	16.0	10.0	4.0	10.0	31.0	10.3
Min	3.0	4.0	6.0	23.0	11.0	2.0	4.0	6.0	18.0	7.0
Max	31.0	32.0	73.0	116.0	38.7	20.0	29.0	31.0	65.0	21.7

B.2.3 Number of clicks and visited pages for Mixed-initiative condition

	Easy									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	4	4	4	12	4	2	2	2	6	2
User 2	2	2	2	6	2	1	1	1	3	1
User 3	4	4	4	12	4	2	2	2	6	2
User 4	5	4	4	13	4	2	2	2	6	2
User 5	4	4	4	12	4	2	2	2	6	2
User 6	4	3	2	9	3	2	2	2	6	2
User 7	4	6	4	14	5	2	2	2	6	2
User 8	4	4	4	12	4	2	2	2	6	2
User 9	4	2	4	10	3	2	1	2	5	2
User 10	5	4	4	13	4	2	2	2	6	2
User 11	4	4	4	12	4	2	2	2	6	2
User 12	4	2	2	8	3	2	1	1	4	1
User 13	2	5	2	9	3	1	2	1	4	1
User 14	4	4	4	12	4	2	2	2	6	2
User 15	4	4	4	12	4	2	2	2	6	2
Average	3.9	3.7	3.5	11.1	3.7	1.9	1.8	1.8	5.5	1.8
Median	4.0	4.0	4.0	12.0	4.0	2.0	2.0	2.0	6.0	2.0
Mode	4.0	4.0	4.0	12.0	4.0	2.0	2.0	2.0	6.0	2.0
Min	2.0	2.0	2.0	6.0	2.0	1.0	1.0	1.0	3.0	1.0
Max	5.0	6.0	4.0	14.0	4.7	2.0	2.0	2.0	6.0	2.0

	Medium									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	10	11	2	23	8	4	2	2	8	3
User 2	5	5	4	14	5	2	2	2	6	2
User 3	9	6	4	19	6	3	2	2	7	2
User 4	11	6	3	20	7	6	2	2	10	3
User 5	12	7	6	25	8	6	2	4	12	4
User 6	3	4	3	10	3	2	2	2	6	2
User 7	6	2	2	10	3	4	2	2	8	3
User 8	2	2	4	8	3	2	2	2	6	2
User 9	5	2	5	12	4	2	2	2	6	2
User 10	6	7	4	17	6	4	2	2	8	3
User 11	9	4	3	16	5	3	2	2	7	2
User 12	10	4	6	20	7	6	2	4	12	4
User 13	12	10	2	24	8	6	2	2	10	3
User 14	2	7	5	14	5	2	2	2	6	2
User 15	11	2	2	15	5	6	2	2	10	3
Average	7.5	5.3	3.7	16.5	5.5	3.9	2.0	2.3	8.1	2.7
Median	9.0	5.0	4.0	16.0	5.3	4.0	2.0	2.0	8.0	2.7
Mode	10.0	2.0	2.0	14.0	4.7	2.0	2.0	2.0	6.0	2.0
Min	2.0	2.0	2.0	8.0	2.7	2.0	2.0	2.0	6.0	2.0
Max	12.0	11.0	6.0	25.0	8.3	6.0	2.0	4.0	12.0	4.0

	Complex									
		Clicks				Pages				
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	10	7	12	29	10	9	5	12	26	9
User 2	12	7	14	33	11	9	5	12	26	9
User 3	46	27	19	92	31	35	24	11	70	23
User 4	15	26	12	53	18	19	23	10	52	17
User 5	8	21	14	43	14	8	17	12	37	12
User 6	7	25	13	45	15	6	22	12	40	13
User 7	17	30	21	68	23	14	26	15	55	18
User 8	15	14	19	48	16	12	11	16	39	13
User 9	17	29	10	56	19	15	23	10	48	16
User 10	15	27	19	61	20	14	24	11	49	16
User 11	12	26	14	52	17	9	23	12	44	15
User 12	8	10	19	37	12	8	10	15	33	11
User 13	41	7	12	60	20	33	5	12	50	17
User 14	7	30	21	58	19	6	20	15	41	14
User 15	12	14	14	40	13	9	11	12	32	11
Average	16.13	20.00	15.53	51.67	17.22	13.73	16.60	12.47	42.80	14.27
Median	12	25	14	52	17.3	9	20	12	41	13.6
Mode	12	7	14			9	5	12	26	8.6
Min	7	7	10	29	9.6	6	5	10	26	8.6
Max	46	30	21	92	30.6	35	26	16	70	23.3

B.2.4 Number of clicks and visited pages for Static condition

	Easy									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	7	3	6	16	5	7	4	7	18	6
User 2	6	8	13	27	9	6	8	12	26	9
User 3	3	3	7	13	4	4	4	8	16	5
User 4	3	8	5	16	5	3	3	3	9	3
User 5	3	7	7	17	6	3	3	5	11	4
User 6	6	4	7	17	6	3	3	5	11	4
User 7	3	4	7	14	5	3	3	5	11	4
User 8	3	8	7	18	6	4	6	8	18	6
User 9	4	3	7	14	5	3	3	5	11	4
User 10	5	3	7	15	5	6	4	8	18	6
User 11	3	4	8	15	5	3	3	5	11	4
User 12	9	4	5	18	6	5	3	3	11	4
User 13	6	3	8	17	6	4	4	8	16	5
User 14	3	5	7	15	5	4	6	5	15	5
User 15	7	4	9	20	7	5	3	7	15	5
Average	4.7	4.7	7.3	16.8	5.6	4.2	4.0	6.3	14.5	4.8
Median	4.0	4.0	7.0	16.0	5.3	4.0	3.0	5.0	15.0	5.0
Mode	3.0	3.0	7.0	17.0	5.7	3.0	3.0	5.0	11.0	3.7
Min	3.0	3.0	5.0	13.0	4.3	3.0	3.0	3.0	9.0	3.0
Max	9.0	8.0	13.0	27.0	9.0	7.0	8.0	12.0	26.0	8.7

	Medium									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	33	5	7	45	15	34	6	8	48	16
User 2	16	6	6	28	9	15	6	6	27	9
User 3	23	14	14	51	17	24	15	13	52	17
User 4	12	14	15	41	14	6	3	7	16	5
User 5	13	14	14	41	14	9	3	7	19	6
User 6	12	15	14	41	14	7	3	7	17	6
User 7	14	13	20	47	16	13	3	13	29	10
User 8	22	14	17	53	18	9	16	17	42	14
User 9	23	14	16	53	18	23	3	9	35	12
User 10	19	3	1	23	8	7	4	6	17	6
User 11	21	3	10	34	11	9	3	3	15	5
User 12	43	3	3	49	16	32	3	3	38	13
User 13	12	16	22	50	17	13	15	19	47	16
User 14	4	16	17	37	12	4	15	16	35	12
User 15	19	14	27	60	20	13	3	18	34	11
Average	19.1	10.9	13.5	43.5	14.5	14.5	6.7	10.1	31.4	10.5
Median	19.0	14.0	14.0	45.0	15.0	13.0	3.0	8.0	34.0	11.3
Mode	12.0	14.0	14.0	41.0	13.7	9.0	3.0	7.0	17.0	5.7
Min	4.0	3.0	1.0	23.0	7.7	4.0	3.0	3.0	15.0	5.0
Max	43.0	16.0	27.0	60.0	20.0	34.0	16.0	19.0	52.0	17.3

	Complex									
		Clicks					Pages			
User no	No. Clicks	No. Clicks	No. Clicks	Sum	Average	No. Pages	No. Pages	No. Pages	Sum	Average
User 1	7	38	42	87	29	7	39	42	88	29
User 2	14	21	27	62	21	13	14	28	55	18
User 3	12	27	11	50	17	13	28	65	106	35
User 4	14	37	75	126	42	13	32	49	94	31
User 5	87	17	98	202	67	48	15	70	133	44
User 6	11	47	94	152	51	8	43	58	109	36
User 7	34	30	110	174	58	31	28	85	144	48
User 8	11	37	59	107	36	11	36	57	104	35
User 9	23	49	62	134	45	20	45	40	105	35
User 10	7	9	9	25	8	4	9	41	54	18
User 11	17	37	64	118	39	16	34	50	100	33
User 12	7	21	42	70	23	7	14	29	50	17
User 13	29	82	130	241	80	19	39	87	145	48
User 14	10	38	27	75	25	11	38	28	77	26
User 15	20	45	49	114	38	17	39	42	98	33
Average	20.20	35.67	59.93	115.80	38.60	15.87	30.20	51.40	97.47	32.49
Median	14	37	59	114	38	13	34	49	100	33.3
Mode	7	37	42			13	39	42		
Min	7	9	9	25	8.3	4	9	28	50	16.6
Max	87	82	130	241	80.3	48	45	87	145	48.3

Appendix C: Results Obtained from T-test for Main Tasks and Training Tasks of Experiment One

C.1 Efficiency of main tasks (related to section 3.7.4)

C.1.1 Efficiency results obtained from T-test for searching tasks (related to section 3.7.4)

C.1.1.1 T-test of Adaptable vs. Adaptive

t-Test: Two-Sample Assuming Unequal Variances		
	82.11111	*120.4444*
Mean	96.38095	106.3651
Variance	362.3166	718.1357
Observations	14	14
Hypothesized Mean Difference	0	
df	23	
t Stat	-1.1365	
P(T<=t) one-tail	0.133724	
t Critical one-tail	1.713872	
P(T<=t) two-tail	0.267449	
t Critical two-tail	2.068658	

C.1.1.2 T-test of Adaptive vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	82.11111	*133.7778*
Mean	96.38095	183.2371
Variance	362.3166	3627.103
Observations	14	14
Hypothesized Mean Difference	0	
df	16	
t Stat	-5.14529	
P(T<=t) one-tail	4.89E-05	
t Critical one-tail	1.745884	
P(T<=t) two-tail	9.77E-05	
t Critical two-tail	2.119905	

C.1.1.3 T-test of Adaptive vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	82.11111	45.22222
Mean	96.38095	76.72222
Variance	362.3166	85.88557
Observations	14	14
Hypothesized Mean Difference	0	
df	19	
t Stat	3.474421	
P(T<=t) one-tail	0.001269	
t Critical one-tail	1.729133	
P(T<=t) two-tail	0.002538	
t Critical two-tail	2.093024	

C.1.1.4T-test of Adaptable vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	120.4444	*133.7778*
Mean	106.3651	183.2371
Variance	718.1357	3627.103
Observations	14	14
Hypothesized Mean Difference	0	
df	18	
t Stat	-4.3634	
P(T<=t) one-tail	0.000187	
t Critical one-tail	1.734064	
P(T<=t) two-tail	0.000375	
t Critical two-tail	2.100922	

C.1.1.5 T-test of Adaptable vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	120.4444	*45.22222*
Mean	106.3651	76.72222
Variance	718.1357	85.88557
Observations	14	14
Hypothesized Mean Difference	0	
df	16	
t Stat	3.911563	
P(T<=t) one-tail	0.000622	
t Critical one-tail	1.745884	
P(T<=t) two-tail	0.001243	
t Critical two-tail	2.119905	

C.1.1.6 T-test of Mixed-Initiative vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	133.7778	45.22222
Mean	183.2371	76.72222
Variance	3627.103	85.88557
Observations	14	14
Hypothesized Mean Difference	0	
df	14	
t Stat	6.540523	
P(T<=t) one-tail	6.56E-06	
t Critical one-tail	1.76131	
P(T<=t) two-tail	1.31E-05	
t Critical two-tail	2.144787	

C.1.2 Efficiency results obtained from T-test for purchasing tasks (related to section 3.7.4)

C.1.2.1 T-test of Adaptable vs. Adaptive

t-Test: Two-Sample Assuming Unequal Variances		
	43.83333	*20.33333*
Mean	36.9881	29.32143
Variance	119.5105	56.84814
Observations	14	14
Hypothesized Mean Difference	0	
df	23	
t Stat	2.160092	
P(T<=t) one-tail	0.020714	
t Critical one-tail	1.713872	
P(T<=t) two-tail	0.041428	
t Critical two-tail	2.068658	

C.1.2.2 T-test of Adaptive vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	43.83333	*24.03167*
Mean	36.9881	34.80167
Variance	119.5105	46.21605
Observations	14	14
Hypothesized Mean Difference	0	
df	22	
t Stat	0.635482	
P(T<=t) one-tail	0.265835	
t Critical one-tail	1.717144	
P(T<=t) two-tail	0.531669	
t Critical two-tail	2.073873	

C.1.2.3 T-test of Adaptive vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	43.83333	*24.16667*
Mean	36.9881	32.11905
Variance	119.5105	45.50611
Observations	14	14
Hypothesized Mean Difference	0	
df	22	
t Stat	1.418222	
P(T<=t) one-tail	0.085068	
t Critical one-tail	1.717144	
P(T<=t) two-tail	0.170137	
t Critical two-tail	2.073873	

C.1.2.4T-test of Adaptable vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	24.16667	*24.03167*
Mean	32.11905	34.80167
Variance	45.50611	46.21605
Observations	14	14
Hypothesized Mean Difference	0	
df	26	
t Stat	-1.04806	
P(T<=t) one-tail	0.152128	
t Critical one-tail	1.705618	
P(T<=t) two-tail	0.304257	
t Critical two-tail	2.055529	

C.1.2.5 T-test of Adaptable vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	24.16667	*20.33333*
Mean	32.11905	29.32143
Variance	45.50611	56.84814
Observations	14	14
Hypothesized Mean Difference	0	
df	26	
t Stat	1.034665	
P(T<=t) one-tail	0.155177	
t Critical one-tail	1.705618	
P(T<=t) two-tail	0.310354	
t Critical two-tail	2.055529	

C.1.2.6 T-test of Mixed-Initiative vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	24.16667	*24.03167*
Mean	32.11905	34.80167
Variance	45.50611	46.21605
Observations	14	14
Hypothesized Mean Difference	0	
df	26	
t Stat	-1.04806	
P(T<=t) one-tail	0.152128	
t Critical one-tail	1.705618	
P(T<=t) two-tail	0.304257	
t Critical two-tail	2.055529	

C.2 Efficiency of training tasks (related to section 3.7.4)

C.2.1 Efficiency results obtained from T-test for searching tasks (related to section 3.7.4)

C.2.1.1 T-test of Adaptable vs. Adaptive

t-Test: Two-Sample Assuming Unequal Variances		
	31.66667	*44*
Mean	63.5	41.19048
Variance	302.3376	89.31136
Observations	14	14
Hypothesized Mean Difference	0	
df	20	
t Stat	4.217993	
P(T<=t) one-tail	0.000211	
t Critical one-tail	1.724718	
P(T<=t) two-tail	0.000423	
t Critical two-tail	2.085963	

C.2.1.2 T-test of Adaptive vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	31.66667	*40.66667*
Mean	63.5	70.28571
Variance	302.3376	408.9377
Observations	14	14
Hypothesized Mean Difference	0	
df	25	
t Stat	-0.95201	
P(T<=t) one-tail	0.175104	
t Critical one-tail	1.708141	
P(T<=t) two-tail	0.350207	
t Critical two-tail	2.059539	

C.2.1.3 T-test of Adaptive vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	31.66667	*25.66667*
Mean	63.5	39.28571
Variance	302.3376	101.5018
Observations	14	14
Hypothesized Mean Difference	0	
df	21	
t Stat	4.508492	
P(T<=t) one-tail	9.63E-05	
t Critical one-tail	1.720743	
P(T<=t) two-tail	0.000193	
t Critical two-tail	2.079614	

C.2.1.4T-test of Adaptable vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	44	*40.66667*
Mean	41.19048	70.28571
Variance	89.31136	408.9377
Observations	14	14
Hypothesized Mean Difference	0	
df	18	
t Stat	-4.87711	
P(T<=t) one-tail	6.06E-05	
t Critical one-tail	1.734064	
P(T<=t) two-tail	0.000121	
t Critical two-tail	2.100922	

C.2.1.5 T-test of Adaptable vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	44	*25.66667*
Mean	41.19048	39.28571
Variance	89.31136	101.5018
Observations	14	14
Hypothesized Mean Difference	0	
df	26	
t Stat	0.515942	
P(T<=t) one-tail	0.305128	
t Critical one-tail	1.705618	
P(T<=t) two-tail	0.610255	
t Critical two-tail	2.055529	

C.2.1.6 T-test of Mixed-Initiative vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	25.66667	*40.66667*
Mean	39.28571	70.28571
Variance	101.5018	408.9377
Observations	14	14
Hypothesized Mean Difference	0	
df	19	
t Stat	-5.13397	
P(T<=t) one-tail	2.95E-05	
t Critical one-tail	1.729133	
P(T<=t) two-tail	5.9E-05	
t Critical two-tail	2.093024	

C.2.2 Efficiency results obtained from T-test for purchasing tasks (related to section 3.7.4)

C.2.2.1 T-test of Adaptable vs. Adaptive

t-Test: Two-Sample Assuming Unequal Variances		
	83.5	*83.5*
Mean	118.3571	80.35714
Variance	1486.67	114.4011
Observations	14	14
Hypothesized Mean Difference	0	
df	15	
t Stat	3.553385	
P(T<=t) one-tail	0.001445	
t Critical one-tail	1.75305	
P(T<=t) two-tail	0.002889	
t Critical two-tail	2.13145	

C.2.2.2 T-test of Adaptive vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	83.5	*81*
Mean	118.3571	120.9643
Variance	1486.67	749.8255
Observations	14	14
Hypothesized Mean Difference	0	
df	23	
t Stat	-0.20627	
P(T<=t) one-tail	0.419196	
t Critical one-tail	1.713872	
P(T<=t) two-tail	0.838392	
t Critical two-tail	2.068658	

C.2.2.3 T-test of Adaptive vs. Mixed-Initiative

t-Test: Two-Sample Assuming Unequal Variances		
	83.5	*39*
Mean	118.3571	73.82143
Variance	1486.67	440.5618
Observations	14	14
Hypothesized Mean Difference	0	
df	20	
t Stat	3.795818	
P(T<=t) one-tail	0.000567	
t Critical one-tail	1.724718	
P(T<=t) two-tail	0.001134	
t Critical two-tail	2.085963	

C.2.2.4T-test of Adaptable vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	83.5	*81*
Mean	80.35714	120.9643
Variance	114.4011	749.8255
Observations	14	14
Hypothesized Mean Difference	0	
df	17	
t Stat	-5.16836	
P(T<=t) one-tail	3.86E-05	
t Critical one-tail	1.739607	
P(T<=t) two-tail	7.72E-05	
t Critical two-tail	2.109816	

C.2.2.5 T-test of Adaptable vs. Mixed-Initiative

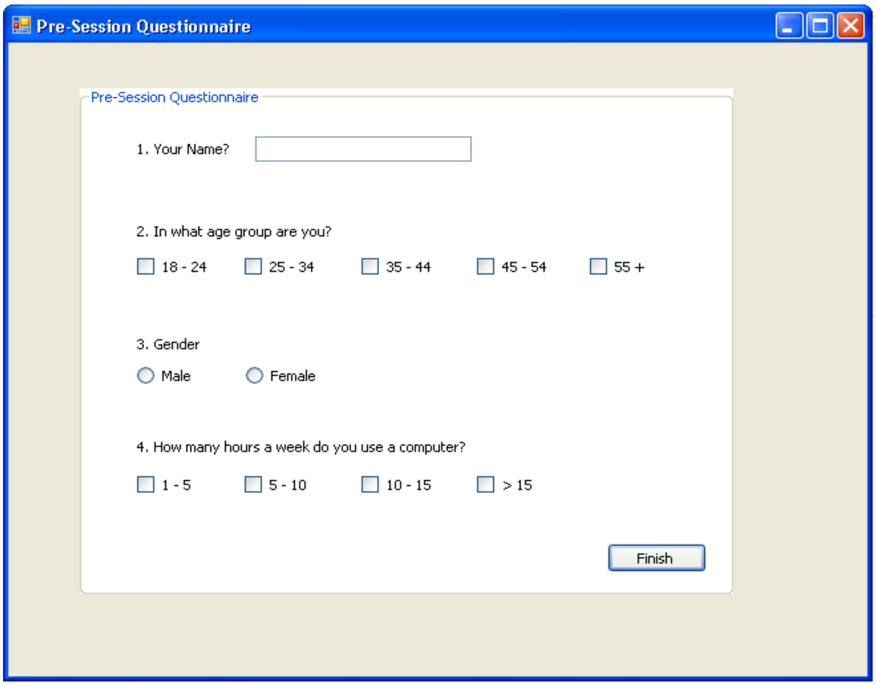
t-Test: Two-Sample Assuming Unequal Variances		
	83.5	*39*
Mean	80.35714	73.82143
Variance	114.4011	440.5618
Observations	14	14
Hypothesized Mean Difference	0	
df	19	
t Stat	1.038066	
P(T<=t) one-tail	0.156135	
t Critical one-tail	1.729133	
P(T<=t) two-tail	0.31227	
t Critical two-tail	2.093024	

C.2.2.6 T-test of Mixed-Initiative vs. Static

t-Test: Two-Sample Assuming Unequal Variances		
	81	*39*
Mean	120.9643	73.82143
Variance	749.8255	440.5618
Observations	14	14
Hypothesized Mean Difference	0	
df	24	
t Stat	5.112529	
P(T<=t) one-tail	1.56E-05	
t Critical one-tail	1.710882	
P(T<=t) two-tail	3.13E-05	
t Critical two-tail	2.063899	

Appendix D: Questionnaire for Experiment Two

D.1 Pre-questionnaire


Pre-Session Questionnaire

Pre-Session Questionnaire

1. Your Name?

2. In what age group are you?

☐ 18 - 24 ☐ 25 - 34 ☐ 35 - 44 ☐ 45 - 54 ☐ 55 +

3. Gender

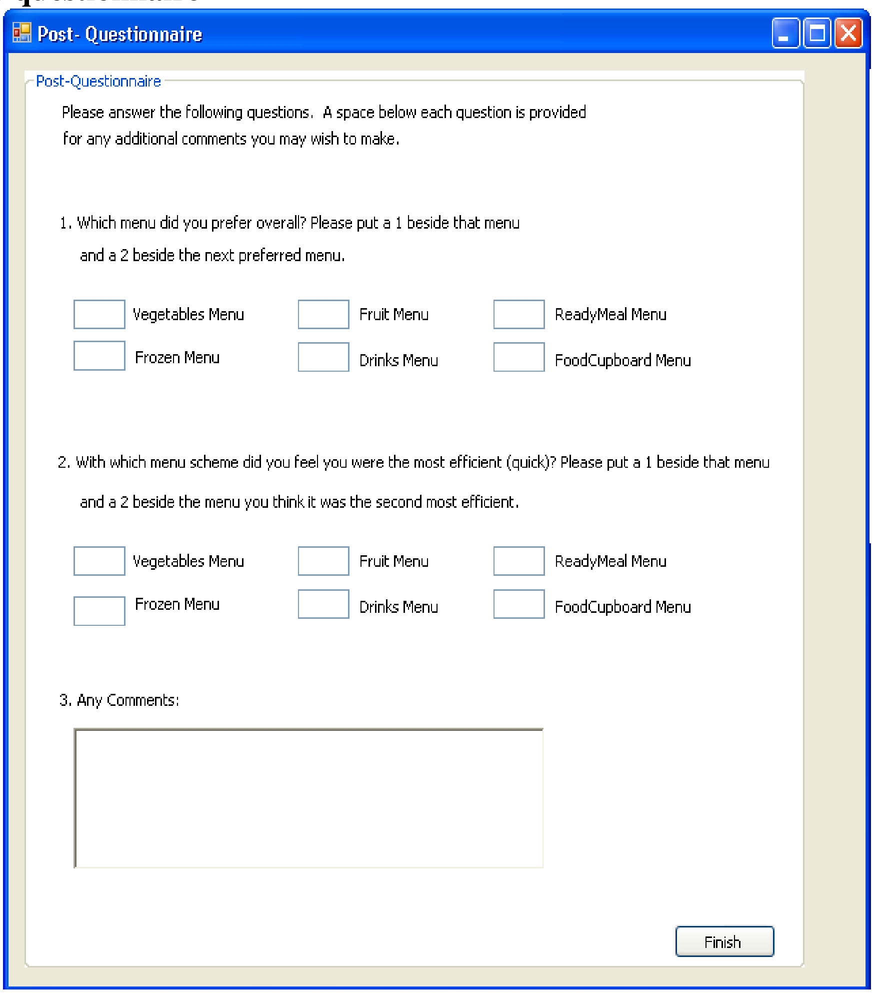
◯ Male ◯ Female

4. How many hours a week do you use a computer?

☐ 1 - 5 ☐ 5 - 10 ☐ 10 - 15 ☐ > 15

Finish

D.2 Post-questionnaire


Post- Questionnaire

Post-Questionnaire

Please answer the following questions. A space below each question is provided for any additional comments you may wish to make.

1. Which menu did you prefer overall? Please put a 1 beside that menu and a 2 beside the next preferred menu.

Vegetables Menu | Fruit Menu | ReadyMeal Menu

Frozen Menu | Drinks Menu | FoodCupboard Menu

2. With which menu scheme did you feel you were the most efficient (quick)? Please put a 1 beside that menu and a 2 beside the menu you think it was the second most efficient.

Vegetables Menu | Fruit Menu | ReadyMeal Menu

Frozen Menu | Drinks Menu | FoodCupboard Menu

3. Any Comments:

Finish

Appendix E: Row Data from Experiment Two

E.1 Task accomplishment time (related to section 4.7.1)

E.1.1 Task accomplishment time for small menus in session 1(related to section 4.7.1.1)

Users	Mixed	Adaptable	Split	Highlighted	Minimised
1	3249	2906	4611	3126	3311
2	2936	2662	2835	2991	2998
3	6256	5493	4485	5085	5736
4	3926	3668	4072	2932	3553
5	5742	4044	4052	4034	4892
6	5338	3914	4414	4661	5148
7	4986	5238	4732	5269	4692
8	4951	4315	5560	5317	5926
9	3281	3125	3462	3577	3142
10	4899	4167	4389	4559	4759
11	3309	2444	3359	2690	3545
12	4896	3887	4492	4138	4976
13	2629	2922	2779	2504	3061
14	4147	2692	3683	3655	3878
15	3426	4107	3682	3433	4794
16	5476	3299	4316	4505	4665
17	2984	3844	3360	3129	3164
18	2816	3068	3471	2728	2945
19	4006	3664	4601	3952	4027
20	2716	2450	3241	2774	2596
21	4100	3951	3354	4121	4964
22	2435	1947	3078	2467	2501
23	3347	4024	4715	3717	3565
24	2259	1962	2402	2451	2276
25	3101	3125	3723	3191	4172
26	2391	3041	2420	3279	2414
27	2158	1728	2361	1896	2151
28	2887	2427	3146	2515	2497
29	2510	2513	2812	2364	2702
30	3200	3401	3833	3867	3566
Average	3678	3334	3715	3498	3754

E.1.2 Task accomplishment time for small menus in session 2(related to section 4.7.1.1)

Users	Mixed	Adaptable	Split	Highlighted	Minimised
1	2657	1930	2748	2265	2247
2	2686	1862	2927	2067	2429
3	3904	3059	4367	3564	4929
4	3764	3079	4000	2493	3843
5	4770	3197	4236	3086	3633
6	3463	3628	4117	4789	4017
7	4102	3031	4244	3770	4692
8	4748	3930	4997	3600	5926
9	3741	2338	2990	2935	3142
10	3819	3192	3789	3740	4759
11	2231	2098	2691	2306	3545
12	4089	3104	4122	3663	4976
13	2344	1679	2142	1837	3061
14	2584	2340	2872	2656	3878
15	3386	2472	3792	3054	4794
16	5028	2986	4305	4120	4665
17	2793	2092	3376	2774	3164
18	2694	2498	2915	2232	2945
19	3192	2558	4448	4071	4027
20	2845	1890	2790	1936	2596
21	2530	2299	2708	2245	4964
22	2136	2013	2849	2165	2501
23	3200	2507	3558	2731	3565
24	2296	1780	2700	1780	2276
25	2908	2587	3395	2123	4172
26	2421	1936	2808	1802	2414
27	2296	1888	2309	1742	2151
28	2805	2517	3024	2130	2497
29	2750	2261	2625	1950	2702
30	3086	2450	3647	2503	3566
Average	3176	2507	3383	2738	3602

E.1.3 Task accomplishment time for large menus in session 1(related to section 4.7.1.2)

Users	Mixed	Adaptable	Split	Highlighted	Minimised
1	3839	4187	4107	4778	4361
2	3454	3929	3456	4216	3680
3	4156	4079	4342	4639	5558
4	4651	3309	4467	4286	4064
5	3605	3814	4009	5461	3870
6	5625	5708	3946	5057	4592
7	5727	4985	4896	4289	5185
8	4258	4000	4045	4677	4901
9	6616	8506	8577	6889	7964
10	5770	5428	6999	5997	5365
11	3742	5706	3209	4004	3022
12	4774	6038	4965	4004	5635
13	4436	3325	3017	3283	5250
14	5054	4979	4472	5368	4766
15	6126	12063	6072	6575	4983
16	6126	4555	4494	4297	5875
17	3479	4752	4672	3208	3979
18	6268	7016	5942	4787	5777
19	5849	6009	7072	7231	5903
20	4100	3951	3354	4121	4964
21	3455	3636	3636	3146	3770
22	3704	4515	4289	4224	4288
23	4156	4872	3500	4174	4143
24	7850	6958	5349	7335	5156
25	4446	4030	3443	3640	4138
26	3905	4283	3372	3732	3665
27	2582	3245	2906	3240	3224
28	2966	3420	3084	3429	3565
29	3151	3649	3121	3467	3982
30	4677	4877	4732	4931	4779
Average	4618	4994	4451	4616	4680

E.1.4 Task accomplishment time for large menus in session 2(related to section 4.7.1.2)

Users	Mixed	Adaptable	Split	Highlighted	Minimised
1	2754	2259	2693	2712	2094
2	2203	2428	2814	3131	1872
3	3554	2683	2904	2582	2274
4	2675	1786	2999	2170	1894
5	2845	2811	2870	2883	2082
6	3297	2520	4268	3665	2516
7	3521	3327	4414	2985	2458
8	3169	3052	3169	2799	2402
9	4087	4987	5816	5057	3978
10	3925	4442	4908	5997	3903
11	2788	3562	2610	2672	2022
12	3859	4289	4567	3332	2557
13	2525	2397	2456	2465	1790
14	2996	4734	3512	3332	2684
15	5693	6793	4977	5544	3607
16	3857	3217	3244	2963	2225
17	3028	1912	3942	2184	2262
18	6027	7014	4356	3298	2931
19	4818	6274	5017	3272	3628
20	2530	2299	2708	2245	2349
21	2483	2047	3036	1860	2069
22	2985	3026	3181	2629	2504
23	3509	3317	2626	5107	3190
24	5064	7253	5359	8427	5015
25	2200	2591	2490	3119	1960
26	2406	1785	4006	1904	2016
27	2349	2341	2533	1952	1384
28	2320	2263	3448	2203	2009
29	2358	2129	3329	2020	1803
30	3165	3078	3588	3332	2499
Average	3300	3421	3595	3261	2533

Appendix F: Results Obtained from T-test for Small and Large Menus of Experiment Two

F.1 Efficiency (related to section 4.7.1)

F.1.1 Efficiency results obtained from T-test for small menus in session 1(related to section 4.7.1.1)

F.1.1.1 T-test of Split vs. Highlighted menus

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3714.65	30	808.197	147.556
	Highlighted	3497.50	30	922.797	168.479

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.807	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	**Std**. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Highlighted	217.144	548.198	100.087	12.444	421.844	2.170	29	.038

F.1.1.2 T-test of Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3714.65	30	808.197	147.556
	Minimise	3753.75	30	1080.079	197.195

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimise	30	.787	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimise	-39.103	667.930	121.947	-288.512	210.307	-.321	29	.751

F.1.1.3 T-test of Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3678.49	30	1149.165	209.808
	Minimise	3753.75	30	1080.079	197.195

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimise	30	.901	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimise	-75.255	500.478	91.374	-262.136	111.627	-.824	29	.417

F.1.1.4 T-test of Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3678.49	30	1149.165	209.808
	Highlighted	3497.50	30	922.797	168.479

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.877	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	180.992	557.862	101.851	-27.317	389.301	1.777	29	.086

F.1.1.5 T-test of Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3334.26	30	908.639	165.894
	Split	3714.65	30	808.197	147.556

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.736	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	-380.388	631.287	115.257	-616.114	-144.661	-3.300	29	.003

F.1.1.6 T-test of Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3334.26	30	908.639	165.894
	Mixed	3678.49	30	1149.165	209.808

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.779	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	-344.236	720.488	131.542	-613.270	-75.201	-2.617	29	.014

F.1.1.7 T-test of Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3334.26	30	908.639	165.894
	Highlighted	3497.50	30	922.797	168.479

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.858	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Highlighted	-163.244	487.884	89.075	-345.423	18.935	-1.833	29	.077

F.1.1.8 T-test of Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3334.26	30	908.639	165.894
	Minimise	3753.75	30	1080.079	197.195

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimise	30	.827	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Minimise	-419.490	608.030	111.011	-646.532	-192.448	-3.779	29	.001

F.1.1.9 T-test of Mixed vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3678.49	30	1149.165	209.808
	Split	3714.65	30	808.197	147.556

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.760	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Mixed - Split	-36.152	749.358	136.813	-315.967	243.663	-.264	29	.793

F.1.1.10 T-test of Minimised vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimise	3753.75	30	1080.079	197.195
	Highlighted	3497.50	30	922.797	168.479

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimise & Highlighted	30	.886	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Minimise - Highlighted	256.246	501.022	91.474	69.162	443.331	2.801	29	.009

F.1.2 Efficiency results obtained from T-test for small menus in session 2(related to section 4.7.1.1)

F. 1.2.1 T-test of Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3383.06	30	741.233	135.330
	Highlighted	2737.67	30	828.936	151.342

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.842	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Split - Highlighted	645.383	449.278	82.027	477.620	813.147	7.868	29	.000

F. 1.2.2 T-test of Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3383.06	30	741.233	135.330
	Minimise	3602.47	30	1028.331	187.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimise	30	.761	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Split - Minimise	-219.413	668.593	122.068	-469.070	30.243	-1.797	29	.083

F. 1.2.3 T-test of Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3175.70	30	805.530	147.069
	Minimise	3602.47	30	1028.331	187.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimise	30	.690	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Mixed - Minimise	-426.767	749.980	136.927	-706.814	-146.720	-3.117	29	.004

F. 1.2.4 T-test of Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3175.70	30	805.530	147.069
	Highlighted	2737.67	30	828.936	151.342

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.750	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	438.030	578.327	105.588	222.079	653.981	4.149	29	.000

F. 1.2.5 T-test of Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2506.70	30	573.179	104.648
	Split	3383.06	30	741.233	135.330

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.878	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	-876.355	363.620	66.388	-1012.133	-740.577	-13.201	29	.000

F. 1.2.6 T-test of Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2506.70	30	573.179	104.648
	Mixed	3175.70	30	805.530	147.069

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.833	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Mixed	-669.002	455.987	83.251	-839.270	-498.734	-8.036	29	.000

F. 1.2.7 T-test of Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2506.70	30	573.179	104.648
	Highlighted	2737.67	30	828.936	151.342

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.803	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Highlighted	-230.972	502.971	91.829	-418.784	-43.160	-2.515	29	.018

F. 1.2.8 T-test of Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2506.70	30	573.179	104.648
	Minimise	3602.47	30	1028.331	187.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimise	30	.775	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Minimise	-1095.769	687.463	125.513	-1352.472	-839.066	-8.730	29	.000

F.1.3 Efficiency results obtained from T-test for large menus in session 1(related to section 4.7.1.2)

F. 1.3.1 T-test of Minimised vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	4680.11	30	1016.131	185.519
	Highlighted	4616.20	30	1187.172	216.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.597	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Minimised - Highlighted	63.911	1000.340	182.636	-309.622	437.444	.350	29	.729

F. 1.3.2 T-test of Mixed vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	4618.16	30	1241.190	226.609
	Split	4451.49	30	1351.533	246.755

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.739	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Split	166.670	942.728	172.118	-185.350	518.691	.968	29	.341

F. 1.3.3 T-test of Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	4451.49	30	1351.533	246.755
	Highlighted	4616.20	30	1187.172	216.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.790	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Highlighted	-164.709	836.827	152.783	-477.185	147.767	-1.078	29	.290

F. 1.3.4 T-test of Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	4451.49	30	1351.533	246.755
	Minimised	4680.11	30	1016.131	185.519

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.793	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	-228.620	824.560	150.543	-536.516	79.276	-1.519	29	.140

F. 1.3.5 T-test of Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	4618.16	30	1241.190	226.609
	Minimised	4680.11	30	1016.131	185.519

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.743	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	-61.949	836.302	152.687	-374.230	250.331	-.406	29	.688

F. 1.3.6 T-test of Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	4618.16	30	1241.190	226.609
	Highlighted	4616.20	30	1187.172	216.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.768	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	1.962	828.501	151.263	-307.406	311.329	.013	29	.990

F. 1.3.7 T-test of Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4994.16	30	1823.433	332.912
	Split	4451.49	30	1351.533	246.755

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.689	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	542.673	1325.213	241.950	47.830	1037.516	2.243	29	.033

F. 1.3.8 T-test of Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4994.16	30	1823.433	332.912
	Mixed	4618.16	30	1241.190	226.609

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.678	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Mixed	376.003	1341.059	244.843	-124.757	876.762	1.536	29	.135

F. 1.3.9 T-test of Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4994.16	30	1823.433	332.912
	Highlighted	4616.20	30	1187.172	216.747

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.678	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Highlighted	377.964	1340.616	244.762	-122.630	878.558	1.544	29	.133

F. 1.3.10 T-test of Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4994.16	30	1823.433	332.912
	Minimised	4680.11	30	1016.131	185.519

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.498	.005

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Minimised	314.053	1584.793	289.342	-277.718	905.825	1.085	29	.287

F.1.4 Efficiency results obtained from T-test for large menus in session 2 (related to section 4.7.1.2)

F. 1.4.1 T-test of Highlighted vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Highlighted	3261.41	30	1443.815	263.603
	Split	3594.62	30	958.621	175.019

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Highlighted & Split	30	.624	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Highlighted - Split	-333.211	1129.989	206.307	-755.155	88.734	-1.615	29	.117

F. 1.4.2 T-test of Minimised vs. Highlighted

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	2532.50	30	796.004	145.330
	Highlighted	3261.41	30	1443.815	263.603

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.906	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Minimised – Highlighted	-728.912	797.774	145.653	-1026.806	-431.018	-5.004	29	.000

F. 1.4.3 T-test of Mixed vs. Split

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3299.72	30	1011.238	184.626
	Split	3594.62	30	958.621	175.019

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.723	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Split	-294.905	734.982	134.189	-569.352	-20.458	-2.198	29	.036

F. 1.4.4 T-test of Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3420.52	30	1600.778	292.261
	Split	3594.62	30	958.621	175.019

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.708	.000

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Split	-174.107	1144.196	208.901	-601.357	253.142	-.833	29	.411

F. 1.4.5 T-test of Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3594.62	30	958.621	175.019
	Highlighted	3261.41	30	1443.815	263.603

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.624	.000

Paired Samples Test									
		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Split - Highlighted	333.211	1129.989	206.307	-88.734	755.155	1.615	29	.117

F. 1.4.6 T-test of Split vs. Minimised

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	3594.62	30	958.621	175.019
	Minimised	2532.50	30	796.004	145.330

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.795	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	1062.123	582.274	106.308	844.698	1279.548	9.991	29	.000

F. 1.4.7 T-test of Mixed vs. Minimised

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3299.72	30	1011.238	184.626
	Minimised	2532.50	30	796.004	145.330

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.769	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	767.218	646.430	118.021	525.837	1008.599	6.501	29	.000

F. 1.4.8 T-test of Mixed vs. Highlighted

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3299.72	30	1011.238	184.626
	Highlighted	3261.41	30	1443.815	263.603

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.645	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	Mixed - Highlighted	38.306	1105.890	201.907	-374.640	451.252	.190	29	.851

F. 1.4.9 T-test of Adaptable vs. Mixed

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3420.52	30	1600.778	292.261
	Mixed	3299.72	30	1011.238	184.626

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.909	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	120.797	800.729	146.192	-178.200	419.795	.826	29	.415

F. 1.4.10 T-test of Adaptable vs. Highlighted

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3420.52	30	1600.778	292.261
	Highlighted	3261.41	30	1443.815	263.603

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.725	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Highlighted	159.103	1137.413	207.662	-265.614	583.820	.766	29	.450

F. 1.4.11 T-test of Adaptable vs. Minimised

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	3420.52	30	1600.778	292.261
	Minimised	2532.50	30	796.004	145.330

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.822	.000

Paired Samples Test

		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	95% Confidence Interval of the Difference Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Minimised	888.015	1049.203	191.557	496.237	1279.794	4.636	29	.000

F.2 Effectiveness (related to section 4.7.3)

F.2.1 Tasks completed by all subjects result obtained from T-test for small menus (related to section 4.7.3)

F.2.1.1 T-test of Adaptable vs. Minimised menus

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.77	30	.728	.133
	Minimised	.33	30	.711	.130

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.422	.020

Paired Samples Test									
		Paired Differences							
			Std.	Std. Error	95% Confidence Interval of the Difference				
		Mean	Deviation	Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Minimised	.433	.774	.141	.144	.722	3.067	29	.005

F.2.1.2 T-test Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.77	30	.728	.133
	Highlighted	.63	30	.718	.131

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.688	.000

Paired Samples Test										
		Paired Differences								
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Highlighted	.133	.571	.104	-.080	.347	1.278	29	.211

F.2.1.3 T-test Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.77	30	.728	.133
	Mixed	.33	30	.711	.130

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.622	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Mixed	.433	.626	.114	.200	.667	3.791	29	.001

F.2.1.4 T-test Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.77	30	.728	.133
	Split	.17	30	.461	.084

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.428	.018

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Split	.600	.675	.123	.348	.852	4.871	29	.000

F.2.1.5 T-test Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.33	30	.711	.130
	Highlighted	.63	30	.718	.131

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.652	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Mixed - Highlighted	-.300	.596	.109	-.523	-.077	-2.757	29	.010

F.2.1.6 T-test Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.33	30	.711	.130
	Minimised	.33	30	.711	.130

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.591	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	.000	.643	.117	-.240	.240	.000	29	1.000

F.2.1.7 T-test Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.17	30	.461	.084
	Minimised	.33	30	.711	.130

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.666	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	-.167	.531	.097	-.365	.031	-1.720	29	.096

F.2.1.8 T-test Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.17	30	.461	.084
	Highlighted	.63	30	.718	.131

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.503	.005

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Highlighted	-.467	.629	.115	-.701	-.232	-4.065	29	.000

F.2.1.9 T-test Minimised vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	.33	30	.711	.130
	Split	.17	30	.461	.084

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Split	30	.666	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Minimised - Split	.167	.531	.097	-.031	.365	1.720	29	.096

F.2.1.10 T-test Minimised vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	.33	30	.711	.130
	Highlighted	.63	30	.718	.131

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.517	.003

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Minimised - Highlighted	-.300	.702	.128	-.562	-.038	-2.340	29	.026

F.2.2 Subjects who completed all tasks result obtained from T-test for small menus (related to section 4.7.3)

F.2.2.1 T-test Minimised vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	.07	30	.254	.046
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.200	.288

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Minimised - Highlighted	-.133	.434	.079	-.295	.029	-1.682	29	.103

F.2.2.2 T-test Mixed vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Split	.30	30	.466	.085

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.400	.028

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Split	-.100	.481	.088	-.279	.079	-1.140	29	.264

F.2.2.3 T-test Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.30	30	.466	.085
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.400	.028

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Highlighted	.100	.481	.088	-.079	.279	1.140	29	.264

F.2.2.4 T-test Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.30	30	.466	.085
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.408	.025

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	.233	.430	.079	.073	.394	2.971	29	.006

F.2.2.5 T-test Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.200	.288

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	.133	.434	.079	-.029	.295	1.682	29	.103

F.2.2.6 T-test Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.792	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	.000	.263	.048	-.098	.098	.000	29	1.000

F.2.2.7 T-test Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Split	.30	30	.466	.085

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.385	.036

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	-.167	.461	.084	-.339	.006	-1.980	29	.057

F.2.2.8 T-test Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Mixed	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.294	.115

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	-.067	.450	.082	-.235	.101	-.812	29	.423

F.2.2.9 T-test Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.539	.002

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Highlighted	-.067	.365	.067	-.203	.070	-1.000	29	.326

F.2.2.10 T-test Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.288	.122

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Minimised	.067	.365	.067	-.070	.203	1.000	29	.326

F.2.3 Tasks completed by all subjects result obtained from T-test for large menus (related to section 4.7.3)

F.2.3.1 T-test Minimised vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	.90	30	.481	.088
	Highlighted	1.00	30	.643	.117

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.558	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Minimised - Highlighted	-.100	.548	.100	-.305	.105	-1.000	29	.326

F.2.3.2 T-test Mixed vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.83	30	.747	.136
	Split	.90	30	.845	.154

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.519	.003

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Mixed - Split	-.067	.785	.143	-.360	.226	-.465	29	.645

F.2.3.3 T-test Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.90	30	.845	.154
	Highlighted	1.00	30	.643	.117

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.444	.014

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Split - Highlighted	-.100	.803	.147	-.400	.200	-.682	29	.501

F.2.3.4 T-test Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.90	30	.845	.154
	Minimised	.90	30	.481	.088

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.569	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	.000	.695	.127	-.259	.259	.000	29	1.000

F.2.3.5 T-test Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.83	30	.747	.136
	Minimised	.90	30	.481	.088

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.528	.003

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	-.067	.640	.117	-.306	.172	-.571	29	.573

F.2.3.6 T-test Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.83	30	.747	.136
	Highlighted	1.00	30	.643	.117

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.718	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	-.167	.531	.097	-.365	.031	-1.720	29	.096

F.2.3.7 T-test Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.83	30	.648	.118
	Split	.90	30	.845	.154

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.599	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	-.067	.691	.126	-.325	.192	-.528	29	.601

F.2.3.8 T-test Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.83	30	.648	.118
	Mixed	.83	30	.747	.136

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.582	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	.000	.643	.117	-.240	.240	.000	29	1.000

F.2.3.9 T-test Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.83	30	.648	.118
	Highlighted	1.00	30	.643	.117

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.579	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Highlighted	-.167	.592	.108	-.388	.054	-1.542	29	.134

F.2.3.10 T-test Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.83	30	.648	.118
	Minimised	.90	30	.481	.088

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.498	.005

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Minimised	-.067	.583	.106	-.284	.151	-.626	29	.536

F.2.4 Subjects who completed all tasks result obtained from T-test for large menus (related to section 4.7.3)

F.2.4.1 T-test Adaptable vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Minimised	30	.288	.122

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Minimised	.067	.365	.067	-.070	.203	1.000	29	.326

F.2.4.2 T-test Adaptable vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Highlighted	30	.539	.002

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Highlighted	-.067	.365	.067	-.203	.070	-1.000	29	.326

F.2.4.3 T-test Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Mixed	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	30	.294	.115

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	-.067	.450	.082	-.235	.101	-.812	29	.423

F.2.4.4 T-test Adaptable vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	.13	30	.346	.063
	Split	.30	30	.466	.085

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Split	30	.385	.036

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Split	-.167	.461	.084	-.339	.006	-1.980	29	.057

F.2.4.5 T-tests Mixed vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Highlighted	30	.792	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Highlighted	.000	.263	.048	-.098	.098	.000	29	1.000

F.2.4.6 T-test Mixed vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Minimised	30	.200	.288

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Minimised	.133	.434	.079	-.029	.295	1.682	29	.103

F.2.4.7 T-test Split vs. Minimised

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.30	30	.466	.085
	Minimised	.07	30	.254	.046

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Minimised	30	.408	.025

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Minimised	.233	.430	.079	.073	.394	2.971	29	.006

F.2.4.8 T-test Split vs. Highlighted

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Split	.30	30	.466	.085
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Split & Highlighted	30	.400	.028

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Split - Highlighted	.100	.481	.088	-.079	.279	1.140	29	.264

F.2.4.9 T-test Mixed vs. Split

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	.20	30	.407	.074
	Split	.30	30	.466	.085

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Split	30	.400	.028

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Split	-.100	.481	.088	-.279	.079	-1.140	29	.264

F.2.4.10 T-test Minimised vs. Highlighted

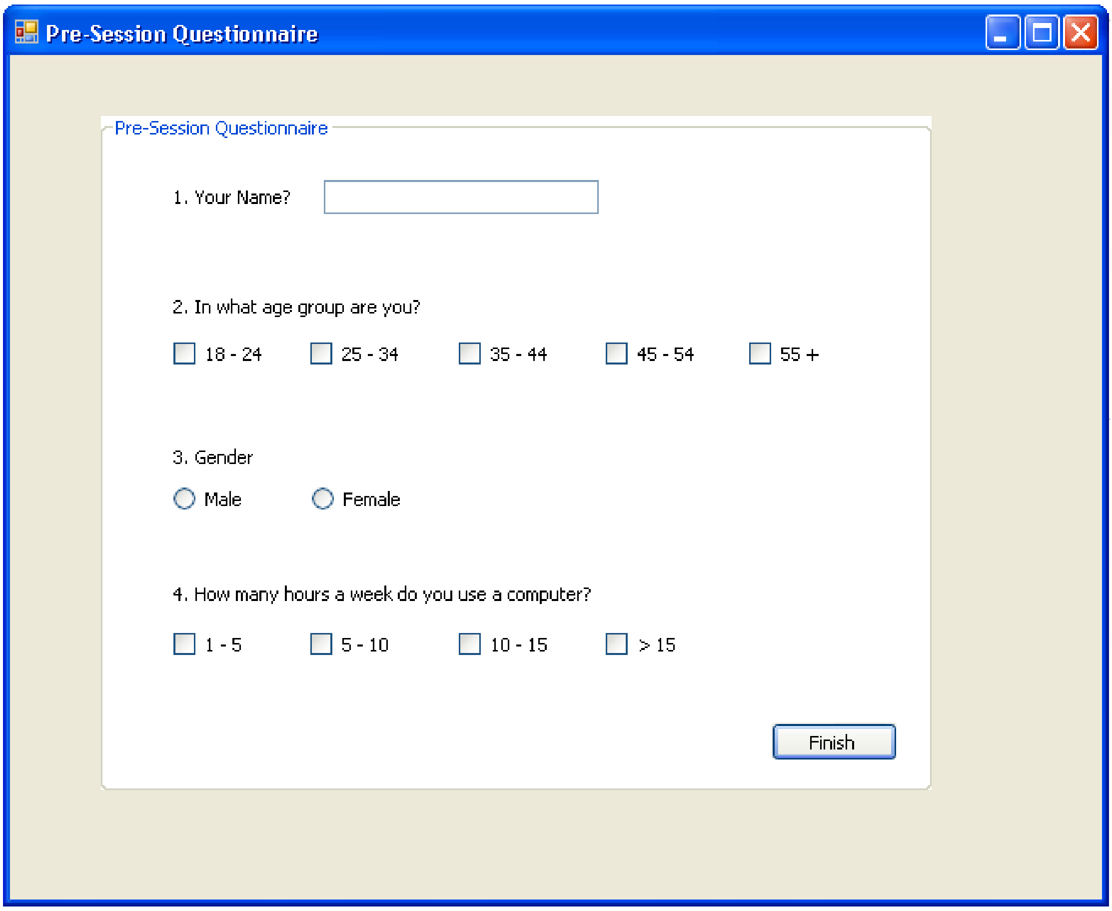
Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Minimised	.07	30	.254	.046
	Highlighted	.20	30	.407	.074

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Minimised & Highlighted	30	.200	.288

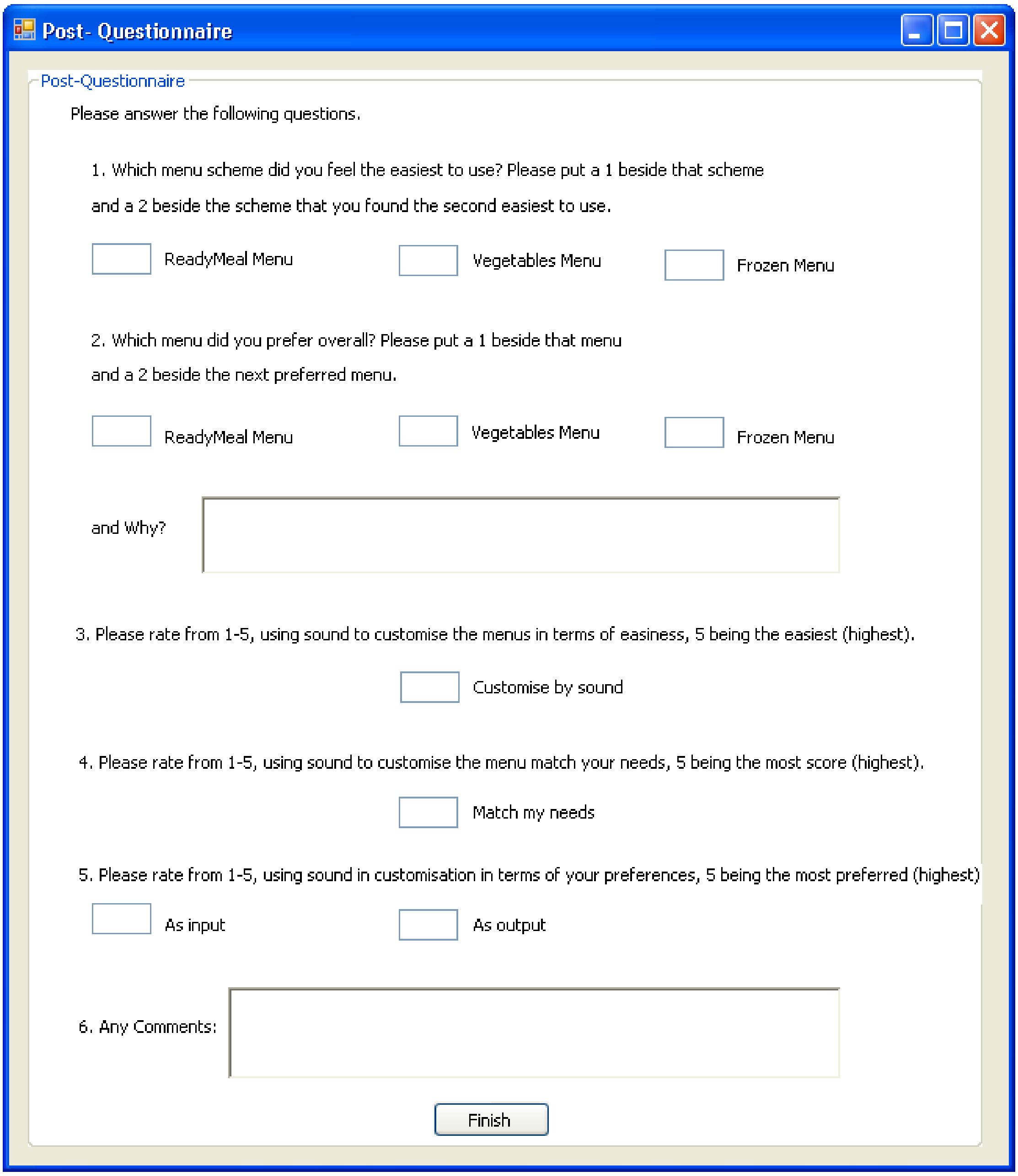
Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Minimised - Highlighted	-.133	.434	.079	-.295	.029	-1.682	29	.103

Appendix G: Questionnaire for Experiment Three

G.1 Pre-Questionnaire



G.2 Post-Questionnaire



Post- Questionnaire

Post-Questionnaire

Please answer the following questions.

1. Which menu scheme did you feel the easiest to use? Please put a 1 beside that scheme and a 2 beside the scheme that you found the second easiest to use.

[ ] ReadyMeal Menu [ ] Vegetables Menu [ ] Frozen Menu

2. Which menu did you prefer overall? Please put a 1 beside that menu and a 2 beside the next preferred menu.

[ ] ReadyMeal Menu [ ] Vegetables Menu [ ] Frozen Menu

and Why?

3. Please rate from 1-5, using sound to customise the menus in terms of easiness, 5 being the easiest (highest).

[ ] Customise by sound

4. Please rate from 1-5, using sound to customise the menu match your needs, 5 being the most score (highest).

[ ] Match my needs

5. Please rate from 1-5, using sound in customisation in terms of your preferences, 5 being the most preferred (highest)

[ ] As input [ ] As output

6. Any Comments:

Finish

Appendix H: Row Data from Experiment Three

H.1 Task accomplishment time (related to section 5.9.1)

H.1.1 Task accomplishment time for small menus from 1 -10 (related to section 5.9.1.1)

Tasks	Users									
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10
T1	10229	12945	7283	10413	10562	4208	4249	5662	4460	4258
T2	2880	6381	5975	4337	3526	2385	2455	5683	3109	2766
T3	2066	1952	3126	2287	2299	10283	2152	5602	2233	3631
T4	4314	13779	1511	1797	1559	3993	2503	5104	3154	2787
T5	3570	1785	2789	3264	1594	3644	3244	5135	2732	4644
T6	3654	1466	1845	9308	2430	6323	2131	5355	2680	3386
T7	2613	2383	4886	10356	6485	3792	2413	6318	4149	3613
T8	3775	1783	5193	4339	6833	4698	2291	4584	3525	2672
T9	1917	1684	3927	4620	2838	5120	1751	2248	1651	2687
T10	2740	2681	4348	3635	4662	2351	2152	2708	2252	2245
T11	4111	2962	4030	5959	2942	3383	2957	4069	2517	3288
T12	2303	1953	1607	1829	1224	3332	3152	4814	3470	2487
T13	6506	2961	1239	1456	1497	2613	1706	5422	4236	3073
T14	2676	5779	3586	2747	3171	4114	1660	4085	3591	2346
T15	1836	1298	2040	1384	1513	3088	1806	6030	4051	3180
T16	2513	3354	3158	3299	6033	3189	1402	3851	2161	2974
T17	6865	2488	3345	2205	2658	3653	1984	5529	3430	2539
T18	2090	1410	1551	1254	1311	4478	1552	6151	2929	3596
T19	7400	3129	4259	1865	3018	5272	1306	4195	2124	4584
T20	3871	5308	2542	2214	2459	2799	1880	2252	3302	1454
T21	1750	1531	3372	3800	1745	1793	1580	2608	4521	1931
T22	1965	2698	1312	3380	2714	3637	2006	5712	1869	2615
T23	1947	2169	1943	1592	2241	4932	2843	4698	3122	3947
T24	1306	2103	1809	1282	1532	2545	1241	6322	2214	1046
T25	2688	3731	9223	3191	3074	2171	2957	4977	2569	2556
T26	1627	1281	2046	1713	2555	9269	1207	5523	1993	1702
T27	1907	1807	3493	4828	28455	4943	4208	4710	3095	2237
T28	3006	2275	4243	4330	11306	2261	3040	3133	5173	2831
T29	6261	2566	3397	2841	5848	3496	3954	4041	3113	3208
T30	1236	1742	1644	1369	2251	5567	1652	5476	2645	1242
T31	1190	1789	1577	2171	1772	8402	1439	6140	3947	1133
T32	2082	11342	3399	5615	3068	2575	1902	3073	4625	2358
T33	1618	3289	2299	5347	2224	3305	1863	2400	1713	1860
T34	2533	2363	5534	1271	1728	4313	4164	4341	2214	3101
T35	1299	5289	1725	2083	2529	4943	3676	4211	2675	2661
T36	1501	1536	1387	999	1852	6529	1512	4891	2072	1130
T37	1240	2241	2134	1465	1741	7538	1264	5157	2254	1027
T38	3737	2642	2143	3016	2274	2396	3646	2193	2100	1633
T39	1278	2105	4067	3121	1812	3527	2834	3894	2623	2917
T40	4410	7249	3036	4987	3671	4537	2102	3438	4140	2552
T41	1213	2210	1566	1251	1152	7122	2115	5816	1773	1082
T42	1194	2258	1750	1411	2087	5025	1552	6670	1664	1010
T43	2091	2274	1117	1487	1447	5196	1405	5567	2234	3495
T44	1273	2673	1328	1544	1352	7327	1551	6253	1797	1361
T45	1359	1431	1562	1545	1473	6416	1591	5735	2003	2102
T46	1561	1954	2116	1999	1952	4061	1336	3667	1483	3285
T47	1833	2275	2169	2319	2040	2717	1394	2409	1804	2097
T48	2036	1264	1312	1058	1049	6214	1512	5904	1854	1184

T49	911	1155	1168	1117	1216	3064	1343	6083	1759	1050
T50	2530	2841	3474	3732	2641	4903	2656	3716	1740	2216
AV	2770	3191	2912	3089	3388	4469	2206	4671	2771	2496

H.1.2 Task accomplishment time for small menus from 11 - 20(related to section 5.9.1.1)

Tasks	**Users**									
	U11	**U12**	**U13**	**U14**	**U15**	**U16**	**U17**	**U18**	**U19**	**U20**
T1	6505	3588	5779	3472	4708	5716	11074	3230	3950	4415
T2	3908	1857	3033	2803	2030	1905	2905	1983	1481	1882
T3	1774	4652	2643	1291	8046	2358	2749	2077	1661	1750
T4	4192	2967	2493	2863	2640	1718	5640	1530	2121	1603
T5	7506	2734	10936	2232	3421	2468	5874	2061	1481	2581
T6	7124	3055	10364	1842	1671	1593	2343	1765	1650	2634
T7	5053	1932	8061	2022	3905	2249	1562	2093	1881	1756
T8	4556	2336	3665	3484	4343	2077	5155	5639	2953	5173
T9	4507	2877	8192	3204	2921	3015	3108	3983	3003	3267
T10	6361	2317	3795	2242	2749	2405	2655	1936	2072	2076
T11	7708	2859	10174	2212	2749	1890	2390	1858	3203	2082
T12	2317	2370	4996	2063	1656	1593	1905	1593	2332	2068
T13	2768	2650	5736	1702	2562	2155	2343	3546	1891	1879
T14	4581	2285	15952	4265	4546	2093	3468	4358	10614	3083
T15	2065	2243	2183	2272	3218	1655	1359	2437	2803	1567
T16	5192	2810	3455	1591	1155	1968	2624	2124	2152	1936
T17	6798	3268	2112	1491	1999	2687	2327	2530	1791	3775
T18	3498	1645	1552	1191	1983	1999	1812	1937	1372	2459
T19	4726	4815	3384	3835	4265	2812	3311	2515	2341	1948
T20	4465	2503	3283	2983	3437	1968	2671	2140	1531	2681
T21	3785	3824	5005	1361	1671	2624	3062	2124	1601	2561
T22	2952	2553	4215	2703	1796	1937	2780	2700	1791	1911
T23	4348	5592	2683	1832	1843	1499	15233	3874	1991	1930
T24	2278	1582	1511	1241	3968	1702	1733	1655	1680	1158
T25	3091	2231	2393	2243	3015	2999	1655	1593	1961	2812
T26	2012	2519	1942	1574	3202	2155	1655	2233	2112	1576
T27	1855	7234	1772	1281	3343	3140	14670	2827	1761	1740
T28	2764	2859	2432	1401	2733	1843	2968	2030	2963	1766
T29	2594	1677	2092	890	983	1452	3577	3827	5546	3610
T30	3519	4571	1912	1231	1608	2046	1562	1530	1901	1343
T31	2292	1492	1740	1391	1718	2296	2421	6374	2843	1258
T32	4992	4218	2733	2513	2812	3952	3249	2874	1721	4170
T33	2304	2285	5717	1151	1468	1796	6608	1655	1731	2984
T34	2773	2117	4135	1845	2187	2499	2655	1624	1621	2030
T35	2641	1786	8752	2002	1140	1687	2202	2749	1341	1605
T36	2358	1469	1271	1311	1421	1077	1577	1546	2502	1667
T37	2813	1698	1521	1732	3046	1702	6202	1780	1391	1459
T38	3094	2085	5958	4174	2233	1859	2296	2671	4745	3083
T39	3155	3397	1090	1972	4905	1687	2546	3171	1611	3102
T40	4418	5990	2302	1862	1234	1796	2655	2390	3223	2256
T41	1974	1596	1522	1421	1640	1187	1437	1296	1660	1394
T42	2632	1515	2252	1351	2202	2187	2077	1999	1771	1669
T43	4040	1859	1711	2343	2546	1468	6749	1921	1921	1308
T44	1716	1382	3574	1141	1296	1155	1640	1265	1290	1372
T45	1675	1385	1571	1762	1937	1921	1796	2077	1410	1480
T46	2823	1572	1982	1181	1499	1546	2765	1812	1720	2729
T47	1852	1840	11455	2082	1280	2718	4202	1905	2604	1578
T48	1455	1216	1501	1612	1655	1062	1530	1327	1481	1062
T49	1581	1262	2012	1101	1515	1328	1843	1390	1771	896
T50	2891	3265	2132	1081	1437	2796	2702	2515	1719	1707
AV	3566	2677	4054	1997	2547	2109	3506	2401	2313	2197

H.1.3 Task accomplishment time for large menus from 1 - 10 (related to section 5.9.1.2)

Tasks	Users									
	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10
T1	7518	9942	8905	6904	5726	5652	4676	7431	7187	6397
T2	9414	4728	4493	5704	10164	6779	2382	5954	4004	3763
T3	2230	6996	2559	9858	3693	5766	2162	6836	4495	4571
T4	5082	4093	3658	6295	3815	1830	2819	8279	4814	5982
T5	4337	5327	3859	8016	5308	5138	2215	11253	5639	5499
T6	2509	9888	4532	6412	8139	7190	3094	8476	4757	4821
T7	6906	3379	4520	7773	8228	10862	7958	9989	5607	5371
T8	10104	8865	5811	6245	5163	5174	3946	7848	5454	2881
T9	3245	6370	3243	4627	2802	1843	7587	4117	3322	2180
T10	5113	12972	3889	5418	5965	9640	4478	5046	4711	3967
T11	7899	4217	4838	5707	6143	3331	2151	6600	4040	3931
T12	4852	2120	2724	11517	11426	5230	1262	1910	5743	5025
T13	5411	3457	13578	10024	10983	1984	1360	2026	4305	8423
T14	15221	5764	2468	3339	3030	3258	7927	7155	8402	7121
T15	1994	3161	10007	9891	9037	2824	1897	2307	3782	3294
T16	7567	3761	6380	7896	4272	5229	7102	6098	3198	2970
T17	5411	6515	4022	9448	11968	7859	1386	1737	5502	3535
T18	3341	5159	9672	1511	7082	2772	973	1598	2292	3027
T19	5808	4144	6966	11478	7449	7437	4354	8521	3611	3302
T20	8328	4863	2112	2449	2823	4646	6792	3224	3451	2461
T21	4113	4739	2838	2218	10457	5117	5370	4300	3184	3002
T22	1536	10657	1780	2000	1599	1729	2366	1699	2410	1472
T23	3065	2440	11421	1501	1978	14234	1701	1593	4537	1904
T24	7805	4601	1678	1698	1184	1380	1632	1492	1874	1659
T25	4254	5865	2831	4456	7920	6527	6777	7602	3903	3583
T26	4011	2944	1549	1856	2458	4361	1972	3372	2662	1794
T27	3327	19122	5749	10334	5136	2828	1150	6816	4236	5483
T28	1722	5976	4560	4954	6042	9858	1491	6094	3358	1595
T29	10264	5190	7014	2020	10203	1212	1312	7582	7212	4903
T30	4091	2183	8538	1630	1181	2435	1816	2185	2920	2226
T31	1743	5945	2085	2832	1648	6459	2969	1894	3032	1667
T32	3631	5595	4371	8776	3705	4397	2245	4977	3893	1770
T33	2115	2130	4944	10321	1896	3023	1435	7933	1902	8800
T34	2169	5833	1900	1488	1351	1860	2049	2126	2129	1353
T35	2070	10520	7619	2563	2496	3262	1201	8494	2314	4085
T36	3659	2115	1501	1107	1113	1259	2138	2288	2041	2774
T37	1513	3864	1661	3651	1640	6782	2298	1626	1861	1605
T38	8608	6281	2361	2026	13633	4698	2242	3339	5032	2256
T39	1560	4456	13971	2351	1849	7367	2613	9072	2322	2220
T40	2106	7116	4988	4580	1177	3209	1435	17885	1550	1758
T41	2237	1558	3430	1070	1171	1059	1893	1948	2537	1518
T42	4535	1515	1608	1697	2187	13455	2644	1779	2862	1655
T43	1258	5445	1711	1312	1449	1422	2227	1809	3384	3427
T44	1599	1958	1462	1046	1351	1168	1322	1732	2356	1628
T45	2104	1208	1503	1278	1272	3305	2830	2013	1739	1969
T46	1566	3280	1877	1916	1737	1033	1171	2015	1579	2020
T47	2130	2017	3441	3074	1600	3456	1165	6427	2281	4371
T48	1350	1513	1606	1167	1023	1710	1614	2215	2248	2192
T49	1270	1514	1241	955	977	1680	1051	1806	1723	1802
T50	1249	3530	5357	9069	4273	2570	1448	7353	2480	1554
AV	4299	5137	4497	4709	4578	4546	2802	4957	3598	3331

H.1.4 Task accomplishment time for large menus from 11 - 20 (related to section 5.9.1.2)

Task	Users									
	U11	U12	U13	U14	U15	U16	U17	U18	U19	U20
T1	9613	8780	10002	5584	12323	6777	11261	5239	2507	5226
T2	9648	2608	22431	4275	3624	3609	4530	1874	3463	2229
T3	2688	2163	12678	1762	2343	5327	2327	5608	2222	1713
T4	3598	8921	20599	2523	2640	8858	7186	9933	6337	1916
T5	5706	3920	13819	10895	5249	5624	6514	2108	3373	1827
T6	7409	2642	11005	3785	1812	4905	6108	2624	2381	2432
T7	24325	2569	8531	2923	9421	6061	6577	3233	2672	2308
T8	11701	7621	10144	12187	11311	6577	6374	7436	2782	1992
T9	6195	2973	11326	4646	18436	8827	17827	12217	23632	4088
T10	4859	9009	7390	6408	10639	4280	7077	9702	4535	3608
T11	9839	10237	11766	4586	2437	4577	7249	2499	2832	5312
T12	1751	3251	1662	1732	12467	6171	3562	3249	2612	1110
T13	4726	4720	3124	2403	13155	2359	4187	2640	19877	4509
T14	10240	3331	16974	4976	4046	3358	4124	6358	5406	5227
T15	2381	2237	9362	1261	8358	1999	1859	3983	2432	1921
T16	3655	5067	3274	7280	2062	3718	6968	4921	17432	2080
T17	3295	2128	1801	5768	1858	2984	11217	4468	1901	2022
T18	1580	1957	1521	1353	4061	4479	2483	2858	2101	1236
T19	8442	2785	8461	5818	1749	3655	8499	1827	9452	2792
T20	8647	5333	6739	4796	3734	1627	8718	6671	2462	5604
T21	5139	2553	2733	5978	1437	4265	31091	6249	4715	1832
T22	2926	2258	1992	1700	2327	2671	6077	4421	2786	2057
T23	5849	2967	1692	1651	1968	1312	2562	2171	2422	1323
T24	3090	1525	1672	2021	1874	1546	2296	1858	2533	1012
T25	11194	2380	2552	2984	2733	3921	5358	2530	2482	2440
T26	2672	1695	2142	1892	2452	1952	2405	2046	1591	1909
T27	4580	3211	17836	5808	1546	3265	3546	3514	2040	2855
T28	4468	6781	6288	2002	1671	6046	2249	1999	2179	1699
T29	9065	13904	4806	1591	1187	1734	1468	1608	8772	1335
T30	5776	2318	1551	1972	2733	2765	1671	1811	1841	1163
T31	2159	1465	1741	1962	1859	1640	3546	1546	2172	1391
T32	8978	3906	7039	5738	3640	7233	2968	3749	4304	3590
T33	6416	4020	4386	1822	1312	1609	2562	5483	2342	3688
T34	4802	2124	2923	1882	3062	4030	4827	2858	2793	1197
T35	3814	1798	9082	2363	1015	2124	2202	2062	4996	1437
T36	2759	1869	1450	2272	2343	4296	1374	2640	1921	1177
T37	4462	1390	1531	2443	1827	1671	5233	1733	3073	1761
T38	4262	3919	3954	3304	8624	3296	4202	4327	12927	2410
T39	5843	1298	4265	2022	1202	2452	2405	1843	4675	1565
T40	14937	8264	5808	3454	1890	3999	4733	4421	3383	1259
T41	3048	2341	1401	2683	1546	2202	2640	1718	1481	1055
T42	1922	1749	1191	2224	1546	2609	2843	1687	2772	1406
T43	2743	6558	2172	2002	2280	2452	2765	3140	2159	1425
T44	2632	2127	1230	2462	1874	2640	1952	1687	3453	998
T45	1933	1444	2162	1772	2405	1859	1671	2390	1470	1259
T46	7319	1684	9055	1551	1390	2265	2187	1859	2172	1797
T47	2703	2425	2853	1311	1187	4983	1655	5780	1461	1729
T48	2434	1811	1592	1722	1733	3546	2343	1921	1441	1101
T49	2921	967	1902	1302	2499	3109	1874	1531	1201	921
T50	5077	2172	1672	2543	2655	2593	3234	1640	5316	1419
AV	5724	3704	6066	3388	3951	3717	5012	3633	4346	2187

Appendix I : Results Obtained from T-test for Small and Large Menus of Experiment Three

I.1 Efficiency (related to section 5.9.1)

I.1.1 Efficiency results obtained from T-test for small menus (related to section 5.9.1.1)

I.1.1.1 T-test of Adaptive vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptive	3124.14	20	637.659	142.585
	Mixed	2911.49	20	604.315	135.129

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptive & Mixed	20	.675	.001

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptive - Mixed	212.654	501.618	112.165	-22.111	447.419	1.896	19	.073

I.1.1.2 T-test of Adaptable vs. Mixed

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2966.42	20	774.360	173.152
	Mixed	2911.49	20	604.315	135.129

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	20	.736	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	54.931	525.437	117.491	-190.981	300.843	.468	19	.645

I.1.1.3 T-test of Adaptable vs. Adaptive

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	2966.42	20	774.360	173.152
	Adaptive	3124.14	20	637.659	142.585

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Adaptive	20	.787	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Adaptive	-157.723	478.387	106.971	-381.615	66.169	-1.474	19	.157

I.1.2 Efficiency results obtained from T-test for large menus (related to section 5.9.1.2)

I. 1.2.1 T-test of Mixed vs. Adaptive

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Mixed	3920.85	20	1008.326	225.469
	Adaptive	3823.26	20	994.819	222.448

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Mixed & Adaptive	20	.740	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Mixed - Adaptive	97.588	722.802	161.623	-240.694	435.870	.604	19	.553

I.1.2.2 T-test of Adaptable vs. Adaptive

Paired Samples Statistics					
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4209.08	20	951.437	212.748
	Adaptive	3823.26	20	994.819	222.448

Paired Samples Correlations				
		N	Correlation	Sig.
Pair 1	Adaptable & Adaptive	20	.778	.000

Paired Samples Test									
		Paired Differences							
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference		t	df	Sig. (2-tailed)
					Lower	Upper			
Pair 1	Adaptable - Adaptive	385.813	650.280	145.407	81.473	690.153	2.653	19	.016

1.1.2.3 T-test of Adaptable vs. Mixed

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Adaptable	4209.08	20	951.437	212.748
	Mixed	3920.85	20	1008.326	225.469

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Adaptable & Mixed	20	.826	.000

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Adaptable - Mixed	288.225	580.183	129.733	16.691	559.759	2.222	19	.039

List of Published Papers

The following publications have been derived from aspects of the research contained in this thesis.

Journal Papers (Submitted/Pending)

- Two journals paper are going to be submitted

- K. Al-Omar and D. I. Rigas, (2009): "Does the Size of Personalized Menus Affect User Performance?," Journal of Computer Science, vol. 5, pp. 937-947.

Conference Papers

- Khalid Al-Omar, Dimitrios Rigas. (2009): "Static, Adaptive, Adaptable, and Mixed-initiative approaches in ecommerce: Controllability and Customisation", Vouliagmeni Beach, Athens, Greece. In press

- Khalid Al-Omar, Dimitrios Rigas. (2009): "The Effect of Size of Personalised Menus on User Satisfaction", Vouliagmeni Beach, Athens, Greece. In press

- Khalid Al-Omar, Dimitrios Rigas. (2009): "Comparison of Adaptive, Adaptable and Mixed-initiative Menus", In Proceedings of International Conference on CYBERWORLDS, IEEE. Bradford, UK.

- Khalid Al-Omar, Dimitrios Rigas. (2009): "A User Performance Evaluation of Personalised Menus", In Proceedings of the Second International Conference on the Applications of Digital Information and Web Technologies (ICADIWT 2009) IEEE. 2009, p.p 104-109, London, UK, 978-1-4244-4457-1/09.

- Al-Omar K. and Rigas D (2009): "Investigating the usability of using Static, Adaptable, Adaptive, and Mixed-initiative keyboards in e-commerce", In Proceedings of the Saudi International Innovation Conference 2009, p.p, Surry, UK.

- Al-Omar K. and Rigas D. (2008): "An Empirical Study to Investigate the Efficiency of Static, Adaptable, Adaptive, and Mixed-Initiative Environments in E-Commerce", In Proceedings of IADIS International Conference on Interfaces and Human-Computer Interaction, pp. 239-243, 25-27 July 2008, InderScience Publishers. IADIS, Amsterdam, Netherlands , ISBN: 978-972-8924-5.

- Al-Omar K. and Rigas D (2008): "A platform for Investigating Effectiveness for Static, Adaptable, Adaptive, and Mix-Initiative Environments in E-Commerce", In Proceedings of International Conference on E-Business, ICETE ICE-B, pp. 191-196, July 26-29 2008, Porto, Portugal, ISBN: 978-989-8111-5.

- Rigas D. and Al-Omar K. (2008): "A platform for Investigating User Satisfaction for Static, Adaptable, Adaptive, and Mix-Initiative Environments in E-Commerce", In Proceedings of the Saudi International Innovation Conference 2008, p.p. 63-67, Leeds, UK, 978-0-9559241-2-5.

Workshop Papers

- Khalid Al-Omar, and Rigas, D. (2008). "An Empirical Investigation of Different Interactive Environments in e-commerce". In Proceedings of the Ninth Informatics Workshop for Research Students, Bradford, UK, (13 June, 2008), pp. 206-209. ISBN: 978 1 85143 251 6.

- Khalid Al-Omar, and Rigas, D. (2007). "A Platform for Investigating Static, Adaptive, Adaptable, and Mixed-initiative environments in e-commerce". In Proceedings of the Eighth Informatics Workshop for Research Students, Bradford, UK, (28 June, 2007), pp. 184-186. ISBN: 978 1 85143 246 2.