Static, Adaptive, Adaptable, and Mixed-initiative approaches in e-commerce: Controllability and Customisation

KHALID AL-OMAR, DIMITRIOS RIGAS
Department of Computing
University of Bradford
Bradford, West Yorkshire, BD7 1DP
UK
K.Al-omar1@Bradford.ac.uk, D.Rigas@Bradford.ac.uk

Abstract: - This paper describes a comparative empirical study that aimed to investigate usability and controllability of four interactive conditions: Static, Adaptive, Adaptable, and Mixed-initiative. Each of these conditions implemented separately as a web-based e-commerce application. The structure of the platform is similar to many web-based e-commerce platforms, except that users can purchase items by clicks along with using four types of keyboard: QWERTY, QWERTY with Keypad, AZERTY and Alphabetical keyboard. The difference between the four conditions applied to the content, items position on the list, and keyboard type. These environments were tested independently by four separate groups of users. Each groups consisted of 15 users. 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 be having similar level of control comparing to the adaptive conditions.

Key-Words: - Adaptive, Adaptable, Mixed-initiative, E-commerce, Performance, Interactive systems, Controllability

1 Introduction
It is becoming increasingly difficult to ignore the effects of growing number of functions in software applications, which leads to increase the number of menus, icons and toolbars. One of the main effects of this growing is that the interfaces become visually complex and very hard to control, which recognised as a phenomenon called by some researchers creeping featurism [1] and others bloatware [2]. To overcome this problem and reduce their visual complexity, interfaces need to provide easy access to the functions that users actually use.

There are three main approaches to personalisation which are adaptive, adaptable, and mixed-initiative. In the adaptive interfaces layout and content dynamically are changed by system to meet user’s needs, while in the adaptable interfaces provide customisation techniques which permit users to adjust their layout and content to suit their needs. Mixed-initiative interfaces combine these two approaches to provide what best suits the user [3]. On the other hand, these approaches distinguish 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 [4].

There has been spirited debate as to which of these approaches is best [5]. Some researcher put forward the view that users should be provided with easily predictable mechanisms to manage their tools, while the other are of the opinion that users need 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 comparing the level of control subjects obtained while utilising these approaches.

2 Previous works
In recent years, there has been an increasing amount of literature on personalisation. However, an examination of the previous research on personalisation reveals contradictory findings. By way of illustration, in a controlled experiment, 26 subjects were asked to search for names in a telephone directory that users can access through a hierarchy of menus and tested it against a static system. The most interesting finding was that subjects performed faster with the adaptive system, and 69% of subjects prefer the adaptive system. Furthermore, result showed that the adaptive system reduces the search paths for repeated names by 35% in time per selection, and reduce 40% in errors per menu. Another study [6] replicated the previous experiment with a larger number of trails. The results of this study showed that the adaptive system is effective and after using the system for long period of time users did begin to perform better with the static interface. Another study carried out a six-week with a 20 participant field study to evaluate
their two interfaces combined together with the adaptive menus in the commercial word processor Microsoft Word 2000. The two interfaces are a personalised interface containing desired features only and a default interface with all the features only. The first four weeks of the study participants used the adaptable interface, then the remaining for the adaptive interface. 65% of participants prefer the adaptable interface and 15% favouring the adaptive interface. The remaining 20% favouring the MsWord 2000 interface. For example, Jameson and Schwarzkopf directly compared automatic recommendations controlled updating of recommendations and a condition where no recommendations were available. The comparison was concerned with content rather than the graphical user interface. In the automatic recommendation (that is, adaptive) system, the updating was performed automatically by the system, while in the controlled updating of recommendations (that is, adaptable) system, 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 difference in performance score between the three conditions. Recently, another study examined a new adaptive technique called ephemeral adaptation. Ephemeral menus recognise predicted items immediately, while remaining items gradually fade in [7]. 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 preferable to static menus.

Most studies in the field of personalisation have only focussed on studying the differences and similarity between the adaptive and adaptable approach. Consequently, there has been a small amount of research into mixed-initiative interfaces, including a study compared directly between the static, adaptive, adaptable, and mixed-initiative, as to which is best in terms of efficiency [8], effectiveness [9], and user satisfaction [10] in e-commerce. They found that subjects were faster in the mixed-initiative condition, followed by the adaptable and adaptive conditions. Subjects were slower in the static condition. Furthermore, 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. Another study which compared an adaptive bar (mixed-initiative system) with the built-in toolbar present in MSWord (adaptable system) [11]. It found that the mixed-initiative system significantly improved performance in one of two experimental tasks. In another study, Burnt et al. [12] 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.

3 Experiment Platform

The experimental platform is a typical web based e-commerce application. The experimental platform utilised four types of interaction conditions: static, adaptable, adaptive and mixed-initiative approach (See Figure 1). Each condition implemented separately. Furthermore, the four interaction conditions applied principally in the contents (Items list), keyboards and layout.

3.1 Contents (Items list)

Items in the main page categories to six groups. Each category consists of 10 to 50 items. Each group displayed the same amount of information. More specifically, items name, items id, items pictures, items price in all category were displayed. The default number of items displayed at initial of the experiment is four items in each group. The other items are hidden and subjects need to search for the required item inside each category. Groups in the static conditions do not change during subjects use. In contrast, in the adaptive conditions after each selection the selected item will move to the top of the list. Then, the lists will counts how many times each items has been used accept the first item and updates the list. On the other hand, in the adaptable and mixed-initiative conditions, subjects allowed to add new lists to the main page and delete an existing list. Also, subjects can change lists positions by drag and drop list from one zone to another. Furthermore, subject can move items to a specific location on the list (Up or down). In addition, subject can customise the number of displayed items (not less than 1 and up to 10 items) for each category. However, in the mixed-initiative approach, subjects can locked and unlocked the list from moving up or down. Also, if subjects attempt to add impersonalise items, list will warn them by displaying a confirm message.
3.2 Keyboard

In order to purchase by keyboard, subject must enter (by mouse clicks) the item Id. However, four types of keyboard were developed: QWERTY, QWERTY with Keypad, AZERTY, and Alphabetical keyboard. Each condition integrated different keyboard schemes. In the static condition, the QWERTY keyboard provided as that most familiar keyboard. In the adaptive condition, subject can choose only one keyboard from the four types of keyboard before starting the experiment. In the adaptable and mixed-initiative conditions, the four types of keyboard provided together and subjects can switch from one types to another at any time. In contrast, in the mixed-initiative condition the QWERTY keyboard suggested to subjects as the default keyboard.

3.4 Static platform

The contents, layout and keyboard do not change during the course of use. Our goal was to design the ideal platform to do the required tasks as efficiently as possible. In order to do that, the content will be used according to pre-determined tasks and placed on the main page. Beside, the QWERTY keyboard has been chosen as it’s the standard keyboard that most subjects are familiar with. Thus, the content and the keyboard is the ideal for carrying out the tasks.

3.5 Adaptive platform

In the adaptive condition, subject can choose only one keyboard from the four types of keyboard before starting the experiment. In the adaptable and mixed-initiative conditions, the four types of keyboard changed during subjects usage. Our goal was to design the predictable personalise approach as possible. Therefore, we asked subjects before using the interface to choose the types of keyboard they prefer, and to choose some new contents based
on our scenarios. However, when the participant started, four items displayed as a default in each web part at the home page. Moreover, the items order in the list changed according to subjects usage by using two algorithms: frequently and recently used items.

3.6 Adaptable platform
In the adaptable platform the layout, contents, and keyboard were changed by subjects before and during subjects usage. Subjects can customise the display with as many items as they like (minimum one item). In addition, the system allows the subjects to add a new content to the home page and move items inside the list. Therefore, changing the contents of the home page been completely let to user responsibly.

3.7 Mixed-initiative platform
In the Mixed-Initiative condition the control is shared. Therefore, our goal was to ensure that the control is shared as fairly as possible. The Mixed-Initiative condition algorithm is 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. Items will be moved up to the top of the list when clicked three times, even if the list locked. Initially, when the website is loaded the default content of the home page is personalised. Moreover, organising the list is the user’s responsibility along with locking the lists.

4 Experimental design and tasks
This study was designed to determine the effect of different personalisation approaches in efficiency [8], effectiveness [9], user satisfaction [10], and controllability. Therefore, in this experiment four objectives had to be attained first to fulfil our goals. The first one was to measure precisely the efficiency of each conditions by timing the tasks completion and counting the number of clicks, visited pages, and errors in each conditions. The second objectives was to measure the effectiveness of each conditions 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 perspective about the ease of use, ease of purchasing, ease of navigation, ease of shopping, and overall satisfaction. Furthermore, it aimed to obtain subjects perspective about the level of control provided by each condition, and the level of control needed by each subject. However, this paper mainly focuses on the last objective. For these objectives, the experiment and tasks was design to fit in a forty fife minute session. The experimental platform was tested empirically by four independent groups, consisting of 15 users. All the groups of users were asked to accomplish the same 12 tasks. These tasks were designed with three complexity levels: easy, medium, and difficult. In order to avoid the learning effect, the order of the task complexity was varied between subjects. The number of available items, item position (location) in the list, number of requirements and guidance was considered when designing the tasks, that is, more than three items available within a list that consists of a maximum of 20 items. The items are positioned at the top, middle and at the end of the list. Thus users can find the item even if the list changes. The number of requirements is less than four. The users are guided to the list by providing the name of the list and the subcategory. For the medium tasks, the number of available items is reduced to two items within a list that consists of more than 30 items. The items are positioned at the middle of the list. The number of requirements is more than four and up to six requirements. The users are guided to the list but not the subcategory, so it is the user’s responsibility to search for items in the subcategory. For the difficult tasks the number of available items is one item within a list that consists of more than 40 items. The items are positioned at the middle of the list, to make sure that users can find the item even if the list changed. The number of requirements is more than seven. In the difficult tasks there is no guidance to items, so it is the user’s responsibility to search for items in all lists and all subcategories.

Table 1: Tasks Design

<table>
<thead>
<tr>
<th>Category</th>
<th>Easy Tasks</th>
<th>Medium Tasks</th>
<th>Complex Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of requirements</td>
<td>1-3</td>
<td>4-5</td>
<td>6-10</td>
</tr>
<tr>
<td>Number of clicks required</td>
<td>1 to 2 clicks</td>
<td>5 to 10 clicks</td>
<td>More than 15 clicks</td>
</tr>
<tr>
<td>Number of visited pages required</td>
<td>0 pages</td>
<td>1-2 pages</td>
<td>More than 5 pages</td>
</tr>
<tr>
<td>List size</td>
<td>10 to 15 items</td>
<td>25 to 30 items</td>
<td>35 to 50 items</td>
</tr>
<tr>
<td>Items availability</td>
<td>3 to more</td>
<td>1-5</td>
<td>1</td>
</tr>
</tbody>
</table>

5 Subjects
This environment was tested empirically by four independent groups, each consist of 15 subjects. The sample consists of 60 subjects (forty four males, and sixteen females) from the general population. The experiment was tested empirically by four independent groups. All the groups were asked to
accomplish the same group of tasks (three easy tasks, three medium tasks and three difficult tasks) then a one learnable task before starting each levels of tasks. Each user attended a five minute training session about their environment before doing the requested tasks.

6 Procedure
The experiment procedure was as follows. (1) Before starting the experiment a questionnaire was used to obtain subjects demographics, computer experience, and customisation experience. (2) Subjects were given a 5 minutes tutorial to explain how to use the system and even explain the benefit from the approach used. (3) Before each group of tasks, a scenario provided along with a one practical learnable task to assist subjects to get use too using the approach. Subjects were told to enquire questions if they need, regarding the environment they evaluating and the experiment procedure. (4) At the finish of each session subjects were asked to give ratings for the tested environment. The performance of each user was observed, recorded and noted in an evaluation form. (5) After each group of tasks, a short break been given in between. Then, a questionnaire was used to obtain subjects view about the current tasks and approach. For the adaptable approach, subjects been encouraged to customise and informed that they have the right not to do therefore. Therefore, subjects were encouraged to customise before starting the experiment and allowed to do therefore at any time they need. In addition, instructions for customisation were given and assist provided to subjects when needed. For the adaptive approach, subjects been asked to register with the system before starting the experiment. Therefore, instructions for registration were given and assist provided to subjects when needed. For the Mixed-Initiative approach, subjects been asked to register with the system first and then customise the system after reading the scenario of our experiment.

7 Results

7.1 Controllability
At the finish of each session subjects were asked to give ratings for 1 to10 rating scale for user control and 1 to10 rating scale for website control. Figure 4 demonstrated the difference between the four conditions. The high score for subjects control was more or less 90% for mixed-initiative and adaptable conditions.

Figure 2: Controllability
On the other hand, in terms of website 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.

7.3 Customisation Time
According to Figure 3 subjects who customise the adaptable condition spent four times more minutes than those who customise the mixed-initiative condition. 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 3: Customisation Time

8 Discussion
One of the main usability issues of personalisation is controllability. In this experiment we attempted to answer some questions concern about controllability. For example, how much control (might by percentages) 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. However, the result indicates that providing more control than users required is causing confusion. Adaptable, for example different mixed-initiative is more preferred by subjects. In addition, The experimental results obtained from the both quantitative and qualitative measures along with a self reported and observed data. In addition, an interview conducted with subjects when needed. However, it was noticeable that subjects who participated in the evaluation of the mixed-initiative were more confident than the static, adaptable and adaptive. In addition, 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.

9 Conclusion
This paper has investigated the use of four interaction conditions: Static, Adaptable, Adaptive, and Mixed-initiative. In this investigation, the aim was to assess the level of controllability. This study has found that generally the mixed-initiative condition was the best in terms of controllability. In addition, it shows that surprisingly subjects who utilised the static condition were found to be having similar level of control comparing to the adaptive conditions. This study has thrown up many questions in need of further investigation, such as whether the presence and absence of multimodal metaphors on these approaches would affect the controllability.

References: