A Controllability and Customisation Approach to Personalised Web Content

Dimitrios Rigas and Khalid Al-Omar

Abstract—This paper reports an empirical study which investigated the controllability and customisation of four interactive conditions: static, adaptive, adaptable and mixed-initiative. Each of these conditions was implemented separately as a web-based e-commerce application. These environments were tested independently by four separate groups, each consisting of 15 users. Results show that the mixed-initiative condition was the best in terms of controllability. In addition, surprisingly, subjects who used the adaptive condition were found to have a similar level of control to those using the static condition.

Keywords—Adaptable, Adaptive, Customisation, Mixed-initiative, Personalisation.

I. INTRODUCTION

More than before 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. However, since users have different needs, abilities and usage. Some researcher suggested the use of multimodal [3, 4] and multimedia [5] metaphors (such as speech [6], earcons [7], and audio [8]). 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 [9]) [10]. 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 [11, 12] to the needs of individual users to mitigate their complexity [11], since each user will have different preferences, needs, experience and abilities [10].

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 [10].

There has been a debate as to which is the better way to customise interfaces [13], 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 [11]. The mixed initiative approach therefore uses both system control and user control at the same time.

II. PREVIOUS WORKS

There has been spirited debate as to which of these approaches is best [13]. For example, 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 results of this study showed 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 [14] 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 [15]. 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. 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 [16]. 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 [17].

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 which compared an adaptive bar (mixed-initiative system) with the built-in toolbar present in MSWord (adaptable system) [18]. It found that the mixed-initiative system significantly improved performance in one of two experimental tasks. In another study, Burnt et al. [19] 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.

III. EXPERIMENT PLATFORM

The experimental platform is a typical web based e-commerce application. For example, subjects had to register first to log in, and 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. The experimental platform utilised four types of interaction

![Fig. 2 types of items list](image-url)
IV. 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. They were also able to change list positions by dragging and dropping lists from one position to another, 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.

V. 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. 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 them. Therefore, 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 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.

A. 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 (Figure 2 (a) and 3 (a)).

B. 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 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 (Figure 2 (a) and 3 (a)).

C. 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 [11] techniques were available for users to move items to a specific location (Figure 2 (c)). 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.

D. 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 2 (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.

VI. EXPERIMENTAL DESIGN

The four platforms were tested empirically using a between-subjects design (i.e. each subject participated in only one condition). This design was considered ideal for 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.
VII. 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 1). In the easy tasks, subject searched within a list comprising a 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 1 description of tasks that users perform during the experiment

<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</td>
<td>1-3</td>
<td>4-5</td>
<td>6-10</td>
</tr>
<tr>
<td>requirements</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guidance Type</td>
<td>Subjects directed to main and subcategory</td>
<td>Subjects directed to main category</td>
<td>No guidance</td>
</tr>
<tr>
<td>Number of clicks required</td>
<td>1 to 2</td>
<td>5 to 10</td>
<td>More than 15</td>
</tr>
<tr>
<td>Number of pages visited</td>
<td>None required</td>
<td>Maximum of 2</td>
<td>More than 5</td>
</tr>
<tr>
<td>Number of items in the list</td>
<td>10 to 15</td>
<td>25 to 30</td>
<td>35 to 50</td>
</tr>
<tr>
<td>Number of items available</td>
<td>More than three</td>
<td>Fewer than three</td>
<td>Only one</td>
</tr>
<tr>
<td>Item name specified</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Item position in list</td>
<td>Top, middle &amp; end</td>
<td>Middle</td>
<td>Middle</td>
</tr>
</tbody>
</table>

VIII. SUBJECTS

The sixty subjects from the general population (forty-four males and sixteen females) who completed were divided into four independent groups of fifteen each for the empirical work. Subjects in 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 training 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.

IX. 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. Subjects 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 to subjects when needed. For the mixed-initiative approach, subjects were asked to register with the system first and then to customise it after reading the experimental scenario.

X. TRAINING AND DATA COLLECTION

Each subject attended a five-minute recorded training session about their environment before attempting the tasks.
Further explanation was also provided when needed. On the other hand, quantitative and qualitative data was collected by recording the experiments and from questionnaires, interviews, observations and written notes. 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.

XI. RESULTS

A. Controllability

At the end of each session, subjects were asked to give ratings for 1 to 10 rating scale for user control and 1 to 10 rating scale for website control. Figure 5 demonstrated the difference between the four conditions. The high score for subjects control was more or less 90% for mixed-initiative and adaptable conditions. 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.

B. Customisation

According to Figure 4 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 (t14 = 9.32, p < 0.05, r =0.928). 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. On the other hand, all the 15 participants in the adaptive condition registered with the system according to the instructions provided. 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.

C. Discussion

As the static, adaptive, adaptable and mixed-initiative approaches have different levels of controllability. More specifically, it examined whether the personalised approaches affected controllability and customisation, as it does with other usability parameters (such as effectiveness [20]). Therefore, this experiment was conducted to address some questions concerning controllability. For example, how much control is enough to allow users to do their tasks easily on each level of tasks (easy, medium, and complex)? Therefore, we asked subjects after performing each level of tasks 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. Although, 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 and raised the feeling of controllability on subjects. In the adaptable approach, the majority of subjects (12) did not wish to customise their environment fully. By 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 looked 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. Furthermore, it was apparently noticeable that subjects spent less time in customisation in the mixed-initiative than the adaptable conditions.

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.

XII. CONCLUSION AND FUTURE WORK

The study reported in this paper assessed the level of controllability of four interaction conditions: static, adaptable, adaptive and mixed-initiative. It found that the mixed-initiative condition was the best in terms of user controllability, followed by the adaptable condition. In addition, it showed that providing more control than users required caused confusion and irritation. In conclusion, further investigation is needed to investigate the factors making some of these approaches more controllable in one context than another.

REFERENCES