A PLATFORM FOR INVESTIGATING EFFECTIVENESS FOR STATIC, ADAPTABLE, ADAPTIVE, AND MIXED-INITIATIVE ENVIRONMENTS IN E-COMMERCE

Khalid Al-Omar and Dimitris Rigas

Department of Computing, University of Bradford, Bradford BD7 1DP, U.K.

Keywords: Adaptive, Adaptable, Mixed-Initiative, Static, Usability, Effectiveness, Interactive.

Abstract: This paper introduces an empirical study to investigate the use of four interaction conditions: Static, Adaptable, Adaptive, and Mixed-initiative. The aim of this study is to compare the effectiveness of these four conditions with regard to the number of tasks completed by all users and the number of users who completed all tasks. In order to carry out this comparative investigation, four experimental interfaces were built separately. These environments were tested independently by four separate groups of users, each group consisting of 15 users. The results demonstrated that in the searching tasks the most effective condition was the Mixed-Initiative. In the learnable tasks the most effective condition was the Adaptable condition. In addition, the Static approach was found to be less effective than all other approaches.

1 INTRODUCTION

Today, software application and e-commerce web-based application (Alotaibi and Alzahrani, 2004) is crowded with functions, icons, menus, and toolbars (McGrenere et al., 2007). In addition, the web-based e-commerce application is crowded in both the Graphical User Interface and content. This is a phenomenon called 'Bloatware' or 'featurism creeping' (McGrenere et al., 2007). This phenomenon makes searching for information and products within e-commerce web-based application very complex (Findlater and McGrenere, 2004) (Te'eni and Feldman, 2001). Therefore, personalising the application to users need and preferences is essential and becomes very important (Findlater and McGrenere, 2004), (Fink et al., 1998). Personalisation is a topic of debate between two communities, the Intelligent User Interface community favouring adaptability (Shneiderman, B. and P. Maes, 1997) at the expense of user freedom and Human Computer Interaction community favoured adaptability (Shneiderman, B. and P. Maes, 1997) at the expense of system help. According to McGrenere et al. (2002) there are three potential ways to personalisation: 1) by users and this is called an adaptable approach. 2) by system and this is called An adaptive approach. 3) by both the users

and system and this is called Mixed-initiative approach which is a combination of adaptable and adaptive approach.

Despite the disagreement in the research community, there are multiple direct comparisons between Static, Adaptable, and Adaptive approaches have shown different results. In 1985, the first study of adaptation was reported by Greenberg and Witten. They demonstrated an adaptive interface for a menu-driven application. In their study users were novices on the task and the interface (Greenberg and Witten, 1985). In addition, Greenberg and Witten (1985) built a directory of telephone numbers that users can access through a hierarchy of menus. Their goal is to reduce the number of key-presses buttons. Their approach is to present items at a level in the hierarchy according to the number of selection. Greenberg and Witten tested their system against a static system in a 26-participant experiment. Their results showed that subjects performed faster with the adaptive system, and 69% of subjects prefer the adaptive system. In addition, they found that the adaptive system reduces the search paths for repeated names, reduce 35% in time per selection, and reduce 40% in errors per menu. Trevellyan and Browne (1987) replicated the Greenberg and Witten's experiment with a larger number of trails because they believe after a large of trails subjects

will be familiar with the static and they can memorised the sequence of key-presses. This would reduce the mean time per menu. However, they found 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. This study did not provide a firm conclusion since the total number of subjects in each interface is 4 subjects.

In 1989, Jeffrey Mitchell and Ben Shneiderman (1989) conducted an experiment to compare an adaptive menu that items positions change dynamically according to frequently clicked item, with a static menu. Sixty-three subjects assigned randomly tried both menus and carried out the same 12 tasks in each menu. Their results showed that static menu faster than the adaptive menu at first group of tasks, and no difference in the second group of tasks. That because, subjects in both groups were able to increase their performance significantly. However, Eighty one percent of the subjects preferred the static menu. Another study introduces a system to provide environment for adapting Excel's interface to particular users (Thomas and Krogsæter, 1993). The result showed that an adaptive component which suggests potentially beneficial adaptations to the user could motivate users to adapt their interface. Jameson and Shwarzkopf (2000) conducted a laboratory experiment with 18 participants a direct comparison between automatic recommendations, controlled updating of recommendations, and no recommendations available. Their comparison concerned about the content rather than the Graphical User Interface. Their results showed that there was no difference on performance score between the three conditions.

In 2002 McGerenere et al. conducted 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. This work extends by Findlater and McGrenere (2004) and they compared between the static, adaptable, and adaptive menus. Their result concludes that the static menu was faster than the adaptive menu and the adaptable menu was

not slower than the static menu. In addition, it shows that the adaptable menu was preferable than the static menu and the static was not preferable to adaptive menu. Another study examined how characteristics of the users' tasks and customisation behaviour affect their performance on those tasks (Bunt et al., 2004). The results confirm that users may not always be able to customise efficiently. The results indicate that customisation is beneficial to reduce tasks time if it done right. Also, indicate that the potential for adaptive support to help users to overcome their difficulties.

In 2005, Tsandilas and Shraefel conducted an empirical study that examined the performance of two adaptation techniques that suggest items in adaptive lists. They compared between the baseline where suggested menu items were highlighted and shrinking interface which reduced the font size of non-suggested elements. The results indicate that the Shrinking information was shown to delay the searching of items that had not been suggested by the system. In addition, the accuracy affected the ability of participants to locate items that were correctly suggested by the system. Gajos et al.(2005) comparing two adaptive interfaces: 1) their Split interface, which is most of the calculator's functionality was placed in a two-level menu. 2) Altered Prominence interface, all functionality was available at the top level of the interface. The study showed user preference for the split interface over the non-adaptive baseline. Another experiment compared the learning performance of static versus dynamic media among a 129 students. Their result showed that the dynamic media (animation lessons) has a high learning performance than the static media (textbook lessons) (Holzinger, 2008).

Despite the debate between the two communities, there has been very little work directly comparing to either an adaptive or adaptable approach with the Mixed-Initiative approach through empirical studies. On example of a such a comparison conducted by Debevc et al. (1996). They compared between their adaptive bars with the built-in toolbar present in MSWord. Their results showed that the mixed-initiative system improved significantly the performance in one of two experimental tasks. Bunt et al. (2007) designed and implemented the MICA (Mixed-Initiative Customisation Assistance) system. Their system provides users with an ability to customise their interfaces according to their needs, but also provides them with system-controlled adaptive support. Their results showed that users prefer the mixed-initiative support. Also, it shows that the MICA's

recommendations improve time on task and decrease customisation time.

2 THE EXPERIMENTAL PLATFORM

An experimental e-commerce web-based platform was developed to be used as a basis for this empirical study. The platform provided four types of interaction conditions: Static, Adaptable, Adaptive, and Mixed-initiative. The structure of the platform is similar to many e-commerce web-based platforms. The difference between the four conditions applied to the contents, layout, and item position on the list.

2.1 The Static Platform

The layout, content, and item position on the list does not change during the course of usage. Our goal was to design the ideal platform to do the required tasks as efficiently as possible.

2.2 The Adaptive Platform

The layout, content and item position on the list does change by system during the course of usage. Adaptation helped users to find items by changing content to their preferences. Our goal was to design the most predictable personalised approach as possible.

Figure 1: Adaptive list.

Therefore, the adaptive approach algorithm dynamically determines item position on the list based on the most frequently and recently used items. The two algorithms are used by Microsoft (Findlater and McGrenere, 2004) and suggested by the literature (Findlater and McGrenere, 2004). For our experiment, once the user clicks the items they will move up to the top of the list (See Figure 1).

2.3 The Adaptable Platform

The layout, content and item position on the list is changed by the user during the course of usage. Our goal was to make the customisation process as easy as possible. Therefore, the Coarse-grained and Fine-grained (Findlater and McGrenere, 2004) customisation techniques were utilised by allowing the user to move items to a specific location (See Figure 2). However, the main page provides two choices for the user to choose from. The first choice is an empty page that is left to the user's decision as to which content to add in. The second choice is full content that has already been suggested. This is because some of the early studies suggested a need to examine full-featured interfaces versus reduced interfaces. However, when the participant started, four items were displayed as a default in each web part of the home page. Subjects can increase the number of displayed items as many items as they like and reduce the number of displayed items not less than one item. In addition, subjects can sort the web contents by item name, id and price and the user can also search in different sub-categories. Subjects can add new content to the home page, delete, and move an existing content to different positions.
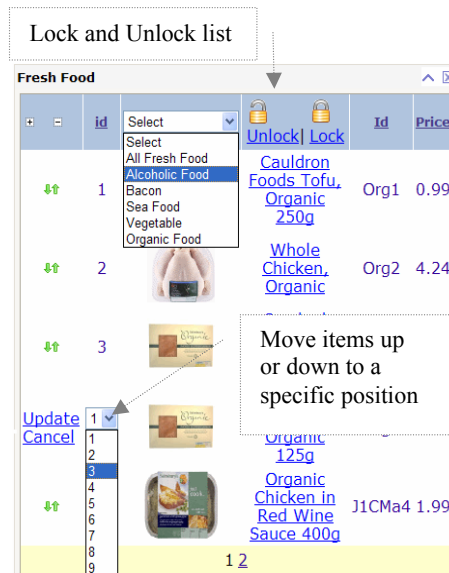

Figure 2: Customisable list.

2.4 The Mixed-Initiative Platform

In the Mixed-Initiative condition the control is shared. Therefore, our goal was to make sure 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 users to take control, a new function was implemented to lock and unlock item movement (See Figure 2). 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. However, organising the list is the user's responsibility along with locking the lists.

3 EXPERIMENTAL DESIGN

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, i.e. 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.

Table 1: Tasks design.

Category	Easy Tasks	Medium Tasks	Complex Tasks
Number of requirements	1-3	4-5	6-10
Number of clicks required	1 to 2 clicks	5 to 10 clicks	More than 15 clicks
Number of visited pages required	0 pages	1-2 pages	More than 5 pages
List size	10 to 15 items	25 to items	35 to 50 items
Items availability	3 to more	1-3	1

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.

4 SUBJECTS

These environments were tested empirically by four independent groups, each group consisting of 15 users. All the groups were asked to accomplish the same group of tasks (three easy tasks, three medium tasks, and three difficult tasks) and a one learnable task before starting each group. Each user attended a five minute training session about their environment before doing the requested tasks. A pre-questionnaire conducted before the experiments to obtain users personal information. All users were between the ages of 18 and 40. 44 of them were male, while the remaining 16 were female. 70% of them were postgraduate students. Most of the participants used the internet for 10 hours or more a week. 85% stated that they do not customise new software unless they have to; the remaining 15 stated that they do so. Also, 32% never used any customisable web pages, where 17% used it once, and just four participants used it every time they went online.

5 RESULTS AND DISCUSSION

Effectiveness was measured by calculating the percentage of users who completed (learning and completion) tasks along with the percentage of (learning and completion) tasks completed by all users. To compare the effectiveness between the four conditions, three critical time limits for task completion was derived for each level of tasks (easy, medium, and complex). Therefore, a task would be regarded as successfully completed if users completed the task within the level critical completion time.

However, it was noticed that during the experiment users who participated in the evaluation
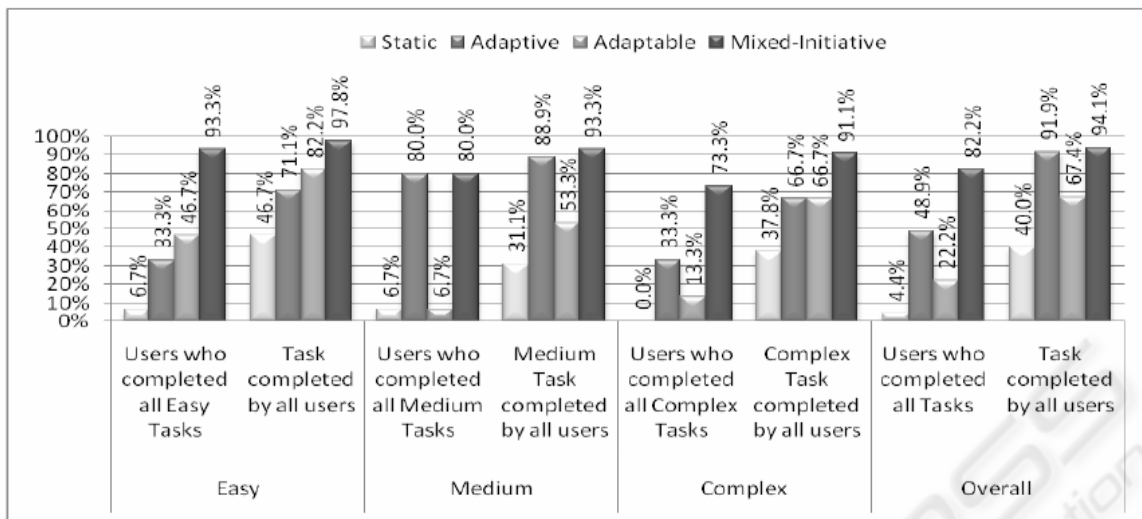
Figure 3: Searching Tasks.

of the adaptable and mixed-initiatives were more confident than the static and adaptive. Also, users got confused while participating in the evaluation of the adaptive and static conditions. This confusion made them spend time on understanding what is happening around them. Overall, just 8 users did not complete all tasks using the Mixed-Initiative whereas 23 users did not complete all tasks using the adaptive condition. In the adaptable condition, 24 users did not complete the all tasks whereas only 2 users did complete all tasks using the Static condition. This shows that the overall number of users who completed all tasks in the Mixed-Initiative is higher than the other conditions. An ANOVA result showed a significant difference in the number of users who completed the tasks at 0.05 (F = (3, 11), p <0.004). The users who completed the easy, medium and complex tasks using the Mixed-initiative condition is higher than the other conditions (Static, Adaptive, and Adaptable), excluding the users who completed the medium tasks using the Adaptive condition.

Overall, t-test was used to find out the difference between the four conditions. t-Test results showed that there was a significant difference of 0.05 between the number of users who completed all tasks using the Mixed-initiative condition compared to the adaptable (t(3)=4.38, cv=3.1) and static (t(3)=11.3, cv=3.1) conditions, but nothing significant was found when compared to adaptive (t(3)=2.04, cv=3.1). The users who completed the tasks using the adaptable and adaptive conditions are higher than the static condition. Also, it was found that the adaptable are higher than the adaptive in easy tasks and lower in medium tasks. Furthermore,

there was a significant difference between the numbers of users who completed all tasks between the adaptable and static conditions (t (3) = 3.04, cv=3.1) and between the adaptive and static conditions (t (4) = 4.5, cv=2.7). Figure 1 shows the percentage of tasks completed by all users in each of the four conditions. However, the number of the tasks completed by all users was calculated to obtain an overall percentage. The result showed that the number of tasks not completed by all users was 8 tasks by using the Mixed-Initiative, 33 tasks by using the Adaptive, 44 tasks by using the Adaptable, and 83 tasks were not completed by using the Static.

In the learnable tasks, there was a difference between the four conditions (See Figure 4). This difference was found to be statistically significant at 0.05 by using the ANOVA test. T-Test results showed that there was a significant difference at 0.05 between the number of tasks completed by all users using the Mixed-initiative condition, compared to the Static condition (t(3)=11.3, cv=3.1) but not to the adaptive (t(2) = 2.6, cv = 4.3) and adaptable conditions (t(2)=3.1, cv=4.3). In addition, there was a significant difference between the Adaptive and Static conditions (t(4) = 4.5, cv=2.7). However, the number of users who completed all learnable tasks by using the adaptable condition was 11, which was higher than the other conditions. Following this was the mixed-initiative where 9 users completed their all learnable tasks, and the Static condition (3 users). The users who completed all tasks using the adaptive condition were lower (2 users) than all other conditions. The percentage of users who completed all tasks using the mixed-initiative condition was higher than the adaptive and static

conditions but not higher than the adaptable condition. The main reason behind this is that sometimes items' positions in the lists changed without users' noticing which caused them confusion.
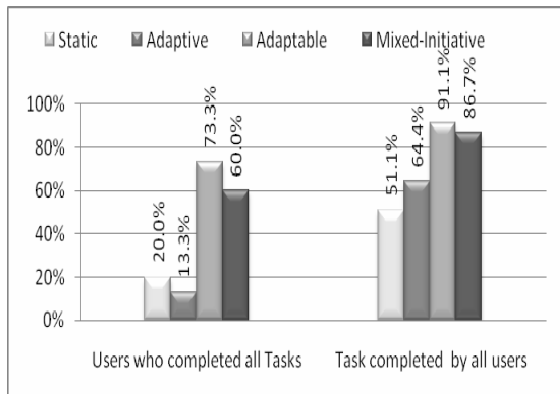

Figure 4: Learnable Tasks.

6 CONCLUSIONS

This paper described an empirical study that was performed to investigate the effectiveness of the Adaptive, Static, Adaptable and Mixed-initiative conditions. In this investigation, the aim was to assess the effectiveness of these four conditions. One of the more significant findings to emerge from this study is that Mixed-Initiative approach was the best in terms of effectiveness in the searching tasks but not with the learnable tasks. In the learnable tasks the adaptable was better than all other approaches. In addition, the Static and adaptive conditions were found to be less effective than the other conditions in terms of number of tasks completed by all users and number of users who completed all tasks. Further work needs to be done to establish whether the presence and absence of multimodal metaphors on the mixed-initiative approach will help to make the most of the adaptive and adaptable advantages, at the same time as reducing their disadvantages.

REFERENCES

Alotabibi, M. & Alzahrani, R. (2004) Evaluating e-Business Adoption: Opportunities and Threats. Journal of King Saud University, 17, 29-41.

Bunt, A., Conati, C. & Mcgrenere, J. (2004) What role can adaptive support play in an adaptable system? Proceedings of the 9th international conference on Intelligent user interfaces, 117-124.

Bunt, A., Conati, C. & Mcgrenere, J. (2007) Supporting interface customization using a mixed-initiative approach. Proceedings of the 12th international conference on Intelligent user interfaces, 92-101.

Debevc, M., Meyer, B., Donlagic, D. & Svecko, R. (1996) Design and evaluation of an adaptive icon toolbar. User Modeling and User-Adapted Interaction, 6, 1-21.

Findlater, L. & Mcgrenere, J. (2004) A comparison of static, adaptive, and adaptable menus. Proceedings of the 2004 conference on Human factors in computing systems, 89-96.

Fink, J., Kobsa, A. & Nill, A. (1998) Adaptable and adaptive information provision for all users, including disabled and elderly people. The New Review of Hypermedia and Multimedia, 4, 163-188.

Gajos, K., Christianson, D., Hoffmann, R., Shaked, T., Henning, K., Long, J. J. & Weld, D. S. (2005) Fast and robust interface generation for ubiquitous applications. Proc. of Ubicomp, 37-55.

Greenberg, S. & Witten, I. H. (1985) Adaptive personalized interfaces—A question of viability. Behaviour & Information Technology, 4, 31-45.

Holzinger, A., Kickmeier-rust, M., & Albert, D. (2008) Dynamic Media in Computer Science Education; Content Complexity and Learning Performance: Is Less More? Educational Technology & Society,, 11, 279-290.

Mcgrenere, J., Baecker, R. M. & Booth, K. S. (2002) 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, 164-170.

Mcgrenere, J., Baecker, R. M. & Booth, K. S. (2007) A field evaluation of an adaptable two-interface design for feature-rich software.

Mitchell, J. & Shneiderman, B. (1989) Dynamic versus static menus: an exploratory comparison. ACM SIGCHI Bulletin, 20, 33-37.

Robert, T. & Dermot, P. B. (1987) A self-regulating adaptive system. SIGCHI Bull., 18, 103-107.

TE'ENI, D. & FELDMAN, R. (2001) Performance and satisfaction in adaptive websites: An experiment on searches within a task-adapted website. Journal of the Association for Information Systems, 2, 1-30.

Thomas, C. G. & Krogsaeter, M. (1993) An adaptive environment for the user interface of Excel. Proceedings of the 1st international conference on Intelligent user interfaces, 123-130.