Comparison of the time-efficiency and perceived workload of different adaptation styles in Self-Adapting Web Menus (SAM)

Margarita Mishinova
Aalto University
margarita.mishinova@aalto.fi

ABSTRACT

Self-Adapting Menus (SAM) were developed to support Internet users in their daily tasks in a web browser through gathering and processing data locally about their behavior and offering adaptation of the menus. Different adaptation styles such as highlighting in yellow and reordering of the menu items support this purpose. They can shorten the visual search time and reduce the mental effort to execute a task. In this study, we examine which of the styles helps the user to find an adapted element on the web menu in an ecological setting with a real webpage and compare it to no adaptation at all. In addition, the perceived mental effort is measured with the NASA-TLX questionnaire because the end user perception has not been studied in detail before. For that, a usability testing was conducted with 14 participants. The "Highlight" style led to faster execution of the task and lower perceived workload. Even though our results are not expandable to the population, we see a clear tendency of the highlighting to improve performance.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI); User interface design;**.

KEYWORDS

Computational design; Adaptive user interfaces, Visual search, Human factors

ACM Reference Format:
Margarita Mishinova. 2018. Comparison of the time-efficiency and perceived workload of different adaptation styles in Self-Adapting Web Menus (SAM). In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Work nowadays consists more and more of higher cognitive and creative tasks and is mostly done on a computer screen and in a web browser. Often the work consists of using a certain website regularly in order to get a task done - for example, an HR specialist using about 30 different job ads platforms to publish the openings of the firm. In such case, the goal of the task is the same on every page, yet the platforms interfaces are organized differently which costs time and mental effort for executing a mundane task. One of the trends in digitization is the automatization of routine working tasks [1]. In the field of Human-computer interaction, Adaptive Web Menus have been developed in order to help users to interact more effectively with a website increasing their efficiency [2].

The open-source SAM [3] is a modular and extensible JavaScript framework for self-adapting menus on web pages. It allows control of which menu items should be adapted (target policy) as well how they are adapted (adaptation style). Through storing and analyzing the user's behavior data locally, it ensures privacy. So far, it consists of six target policies such as Item clicks frequency, or Page visits duration. They build the computational model estimating the importance of the items within a menu. The adaptation styles on the other hand offer different possibilities to show the adapted menu - highlighted in yellow, reordered (with the last used items being shown first), as visible in Figure 1, as well as also combination of these and/or folding the unused menu items. So far, SAM has been tested in laboratory conditions with custom-made websites. Our goal is to explore how is it perceived in an ecological setting with a real web page (in our case, the platform for private announcements, craigslist). We want to find out how SAM influences the performance of the user and also how is it perceived. The research problem is which adaptation style helps the user to find the needed item *n* the menu in the fastest way and how are the two different styles experienced by the end users in an ecologically valid environment.

Figure 1: Craigslist's menu on Housing without adaptation (left) and with the adaptation styles highlighting (middle) and reorder (right).

Our results will help us choose directions for further expanding the capabilities of SAM and making it even more user-friendly.

2 PREVIOUS WORK

Adaptation styles for interfaces have been studied since the 80ies with conflicting results. They have been studied in different contexts such as interface menus[4, 6], software menus[7], and more recently on a website browser [3].

Reorder has proven successful in a study [4] where a hierarchical menu with over a thousand leaf elements was drastically reordered. A possible explanation is that the adaptation reduces the mental effort of the user while the high complexity prevents it from developing strong memory for the location of different elements [5]. In comparison to a static menu, reorder led to a slower task performance and negative user feedback [6]. When using highlighting menu items as a baseline in a study on the accuracy of adaptive algorithms as opposed to shrinking not-suggested menu

items, highlighting led to a faster finding of items among participants [7]. A study [8] on a software calculator interface reports user dissatisfaction with a menu making use of highlight and reordering of items as opposed to a split interface. Wickens also mentions highlighting as possible solution against numerousity and search clutter in maps [9, p. 135] but it can also lead to inference and errors [9, p. 64].

In a recent study on SAM, Gobert et al. (2019) [3] predicted through a Zipfian distribution which items users had to click at each trial while clicking both an adapted item and a non-adapted one. Selection times for adapted elements reduce dramatically while times for selecting non-adapted ones increase. In our current study, we remove the negative influence of selecting non-adapted elements.

3 HUMAN FACTORS

While designing the study, we took following Human factors in consideration. *Visual search* leads to finding a target through shifting the selective attention across a search field. The diameter of the useful field of view (UFOV), the "visual angle within which a target can be detected if it is present, or a non-target identified if it is not" [9], is measured as the separation between the centers of fixation of consecutive eye movements. Since one of the adaptation styles highlights in yellow the already selected items, we hypothesise that executing a certain task in a menu adapted this way will be quicker.

Change blindness is an aspect we should consider when studying the "reorder" adaptation style. It is possible that the change in the environment stays unnoticed for the user. In general, it occurs when the change happens simultaneously with another event, a blink, a blank screen, or outside of the visual field [9]. We assume that users will notice the difference and reordering of the menu items since they will happen inside their visual search field. We expect that the change will increase the effort in executing the exercise so we expect higher perception of the workload in the reorder-group than in the highlight-group.

Another important aspect is *visual clutter* - craigslist as visible in Figure 2 is a good example in which numerous visual elements clutter the interface and slow down the visual search. In order to prepare a rather demanding task for our usability testing, we chose this website.

4 RESEARCH METHOD

In our study, we want to test the usage of SAM on a small scale with end users. The objective of our study is, on the one side, to gather quantitative data on how SAM improves the execution of a task (time in milliseconds) as compared to no adaptation at all, and, on the other side, to learn more about the perception of the adaptation through a standard questionnaire.

Our chosen methods include a between-subjects study design testing the time needed to complete a certain task without an adaptation style and with "Highlight" and "Reorder" style (in randomized order). After completing a task using SAM, the test subjects who will fill out the short version of NASA-TLX, a questionnaire measuring the perceived workload. It includes constructs such as mental demand, effort and frustration which are rated within a 100-points

Figure 2: The homepage of craigslist Finland.

range. The questionnaire is a standard and reliable HCI tool already used in a similar study regarding two differently adapted user interfaces [11].

For cost and time reasons, we choose to make all participants execute a task without adaptation. In order to increase the internal validity, the testing is conducted in a calm, semi-lab-like environment (a study room in the library) and the participants are suitably chosen so that they are experienced computer users of both genders without vision impairments and with good understanding of English.

A threat for the external validity of the study would be the lack of representativeness of the sample for the population - since SAM is not yet open for public use, recruiting a representative sample will not be cost-effective and will slow down the research. The mono-operation bias is avoided through having a control task not using SAM's adaptation styles. The precision of the data is ensured through using a screen capture video recording and a precise tool for measuring the time between two clicks and each testing is recorded and carefully reviewed by the test leader. Learning effects are avoided through recruiting only participants who have never used the website craigslist.com before.

4.1 Hypothesises

The findings of the study will uncover differences in the perception of the two styles, difference in their performance on a before unknown website, and suggest possibilities for further development of the framework. Following hypotheses are examined:

H1. Item selection time will be lower on average in pages with reordered menus than in pages with not-adapted menus. (The user will find the needed item faster through "Reorder" than without an adaptation.)

H2. Item selection time will be lower on average in pages with highlighted menus than in pages with not-adapted menus. (The user will find the needed item faster through "Highlight" than without an adaptation.)

H3. Item selection time will be higher on average on pages with reordered menus than in pages with highlighted menus. (The user will find the needed item faster through "Highlight" than through "Reorder".)

H4. "Reorder"'s NASA-TLX score will be higher on average than "Highlight" (The user will perceive searching an item through "Reorder" as more demanding and effortful than through "Highlight".)

4.2 Study design

The study consists of following parts:

(1) Familiarization phase with no adaptation of the website menu (Learning phase 1)
(2) Executing a certain task on the platform (time is measured) (Task 1)
(3) Familiarization phase with one of the two adaptation styles (Learning phase 2)
(4) Executing a certain task on the platform (time is measured) (Task 2)
(5) Filling out the questionnaire regarding the effort needed for the task execution

For reaching consistent results regarding the time needed to execute a task, each adaptation has its own task in a different menu on craigslist:

(1) No adaptation - finding a certain item in the menu Housing
(2) Highlight - in the menu Services
(3) Reorder - in the menu For sale

Following figures present the tasks in the experiment in the format of a Hierarchical Task Analysis:

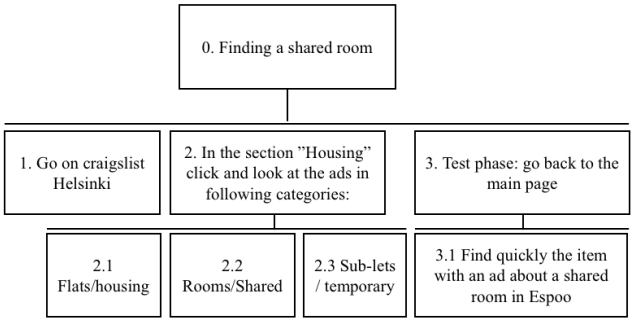

Figure 3: "No adaptation" Task.

4.3 Apparatus and Participants

The experiment run in a room in the university library on a MacBook Air. Participants were asked to sit in a standard chair, in a position they would normally adopt when using a laptop. The screen actions was captured through QuickTime and the time needed for executing the tasks was measured in the aftermath by the experiment lead via an online tool measuring the time between clicks [13]. The tasks were on the website www.helsinki.craigslist.com. The tasks consisted of a longer learning phase in which participants had to click through certain menus on the site and a test phase in which they have to find a certain, already clicked once, category. The

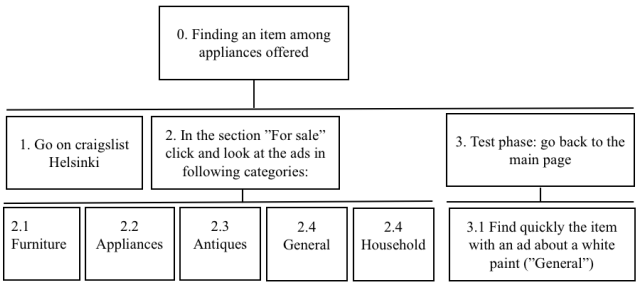

Figure 4: "Highlight" Task.

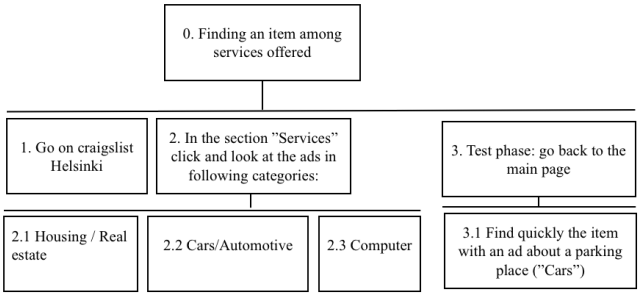

Figure 5: "Reorder" Task.

"No adaptation" task was executed in the Google Chrome browser, Incognito mode, while the "Highlight" and "Reorder" were executed in the Mozilla Firefox browser, also in incognito mode, in order to avoid effects from different users. The target policy was set to Accessrank, an algorithm predicting what item will be chosen next based on recency, frequency, temporal clustering, and time of day [3, 10]. In addition, they filled out the short version of NASA-TLX online [12] .

In the first part of the study, 8 participants without visual impairments took part in the study, 4 of them were female. 4 executed the tasks "No adaptation" and "Highlight" and the other half executed "No adaptation" and "Reorder", the order of the tasks was randomized. After the adapted task, they filled out NASA-TLX only about it. In the second part of the study, 6 more participants (half of them female) tested SAM while this time, they filled out the NASA-TLX questionnaire for both the "No Adaptation" task and the task with adaptation. So we ensured that we also have sufficient data for how the task is perceived without SAM. In the second part of the study, 3 participants executed the tasks "No adaptation" and "Highlight" and the other half executed "No adaptation" and "Reorder". They were recruited in the Computer Science library of the Aalto University. The average age of the sample is 26 years old. 13 have a university education, one is still in high school. None of them had used craigslist before.

5 RESULTS

Data was gathered and processed in Excel (for descriptive statistics and graphs) and R Studio (for inferential statistics).

5.1 Time measures

The item selection time of the three different tasks were measured in milliseconds. Since the tasks differ in the size of the sub-menu in which the users search for the item, we used z-standardization to compare them.

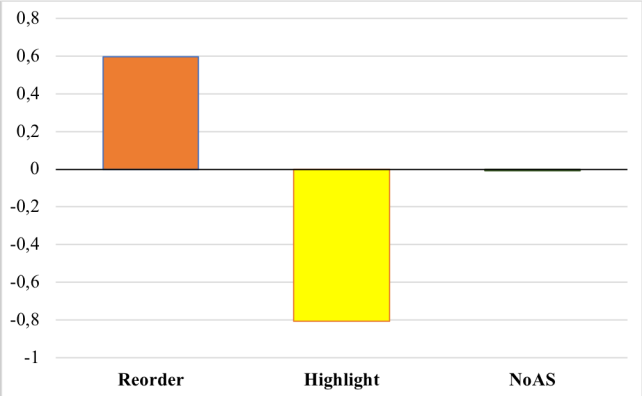

Figure 6: Average scores of the time needed to execute the tasks after z-standardization.

There is a visible tendency that "Reorder" took longer time than "Highlight" or no adaptation at all. We performed an ANOVA to compare the three different styles but the results were not significant even with the bigger sample (f = 0.35445, p = .703212. The result is not significant at $p < .05$). In addition, the different scores were compared with a Mann-Whitney-U-test, a non-parametric test, used when a normal distribution of the data cannot be assumed. The results were not significant:

- "No Adaptation" and "Reorder": U = 192, p = .46414, the result is not significant at $p < .05$;
- "No Adaptation" and "Highlight": U = 194, p = .9681, not significant at $p < .05$;
- "Reorder" and "Highlight": U = 88, p = .65994, not significant at $p < .05$.

Even though, the results cannot lead to a conclusion for the whole population, they give us a valuable insight in how SAM influenced the time needed for executing a task.

5.2 Perceived Workload

The results of NASA-TLX for all the three styles are presented in detail in Figure 7. There is a tendency to perceive "Reorder" as more demanding than "Highlight" and "No Adaptation".

The ANOVA however shows that there is no significant difference between the three scores (f = 0.73982,p = .484724. The result is not significant at $p < .05$). In addition, the different scores were compared with a Mann-Whitney-U-test with following results:

- "No Adaptation" and "Reorder": U = 56,5, p = .58232, the result is not significant at $p < .05$;
- "No Adaptation" and "Highlight": U = 68,5, p = .65994, not significant at $p < .05$;
- "Reorder" and "Highlight": U = 63, p = .29372, not significant at $p < .05$.

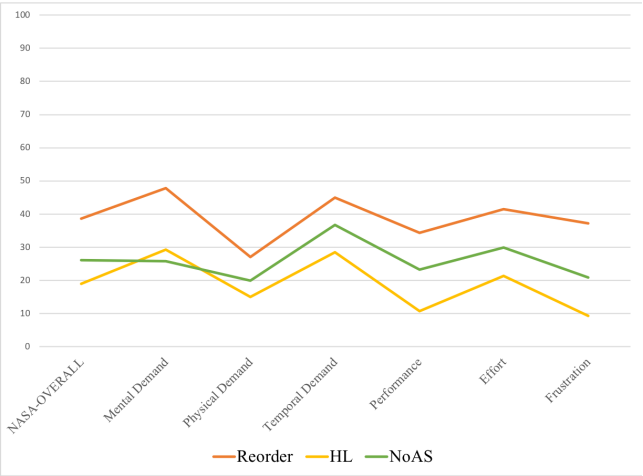

Figure 7: Average scores of NASA-TLX for the three different adaptation styles.

Again, the results cannot lead to a conclusion for the whole population, but they give us a valuable insight in how SAM influenced the perception of the task. Our hypotheses have to be rejected because there is no statistical proof of significance.

6 DISCUSSION

Given the small and unrepresentative sample size, it is no surprise that the results of our study do not show a significant difference between the performance and perception of the two adaptation styles. Nevertheless, they clearly show a tendency for "Highlight" to improve efficiency and lower the mental demand needed for finding a menu item in a new environment. A crucial and unexpected insight was that "Reorder" led to more time needed for executing a task than no adaptation at all.

Another limitation of our study was the shortness of the tasks as well as the methods for measuring the interaction - they give us very brief explanation fo the users behavior. In future studies, longer tasks as well as the usage of eye-tracking can reveal more about SAM's impact.

Our results align with existing findings that adaptation, and especially highlighting, improves the usability of complex menus and are a step forward to wider use of SAM.

7 REFERENCES

[1] McKinsey Global Institute. 2017. A future that works: Automation, Employment, and Productivity.

[2] Oppermann, R. 2017. Adaptive User Support. New York: Routledge.

[3]Gobert, C., Todi, K., Bailly, G. and Oulasvirta, A. 2019. SAM: A Modular Framework for Self-Adapting Web Menus. In *IUI âĂŹ19*, Marina del Ray, CA, USA

[4]Greenberg, S. and Witten, I. H. 1985. Adaptive personalized interfaces: A question of viability. Behaviour and In- formation Technology. 4(1) pp 31-45.

[5] Gajos, K. Z., M. Czerwinski, D. S. Tan, and D. S. Weld. 2006. Exploring the design space for adaptive graphical user interfaces. In *Proceedings of the 2006 Conference on Advanced Visual Interfaces*. 2001-8. New York: ACM.

[6] Mitchell, J. and Shneiderman, B. 1989. Dynamic versus static menus: an exploratory comparison. In *SIGCHI Bull. 20*, 4, 33-37.

[7] Tsandilas, T. and Schraefel, M. C. 2005. An empirical assessment of adaptation techniques. In *CHI '05 Extended Abstracts*.

[8] Gajos, K., Christianson, D., Hoffmann, R., Shaked, T., Henning, K., Long, J.J., and Weld, D.S. 2005. Fast and robust interface generation for ubiquitous applications. In *Ubicomp'05*. Springer.

[9] Wickens, C. D., Parasuraman, R., Banbury, S., and Hollands, J. G. 2015. Engineering Psychology and Human Performance. sychology Press.

[10] Fitchett, S. and Cockburn, A. 2012. Accessrank: Predicting what users will do next. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI âĂŹ12*, p. 2239-2242.

[11] Ahlrichs, Cl. 2011. Development and Evaluation of an Abstract User Interface for Performing Maintenance Scenarios with Wearable Computers. Thesis at the University of Bremen.

[12] NASA-TLX short version online: http://jensgrubert.bplaced.net/nasa-tlx-short/TLX-English-short.html

[13] Time between clicks on same image http://www.yorku.ca/nmw/datt1939f17/javascript$_{a}ll/js_{t}imebetweenclicks.html$

8 APPENDICES

8.1 Consent form

What this study is about

The purpose of this study is to test the efficiency of two different menu adaptation styles. Your participation in this study will help us advance in the research of Self-adaptive menus. *Your participation in this study is voluntary.*

You can take a break at any time. Just tell the researcher if you need a break. You can leave at any time without giving a reason.

Information we want to collect

The data will be used for scientific purposes only and are confidential. All data will be anonymized. No explicit clues of your identity will be left to the stored data. All data will be stored securely. According to the data storage plan of the project, the anonymous data will be made available to other researchers.

How we ensure your privacy

The data will be used for scientific purposes only and are confidential. All data will be anonymized. No explicit clues of your identity will be left to the stored data. All data will be stored securely. According to the data storage plan of the project, the anonymous data will be made available to other researchers. *Your consent*

Please sign this form showing that you consent to us collecting these data. I give my consent:

- For the session to be recorded through Screen capture.

I have received sufficient information about the research study, I have had the possibility to have my questions answered, I have understood the information and I wish to participate in the research study.

Signature and name of participant

If you want to withdraw your consent in the future, contact the person named below who will destroy any personal data we hold about you (such as the recordings).

Contact details: Margarita Mishinova, +49 152 54967 7747, margarita.mishinova@aalto.fi Aalto University

8.2 Demographic questionnaire

Age:
Gender:
Highest educational level:
Have you used craigslist before? Yes no

8.3 Participants instructions

8.3.1 No Adaptation - Task 1.1. Finding a place to stay in Helsinki/Espoo area (learning phase)

Imagine that you have to leave your current flat and find a place to stay from June on. You go on craigslist to search for such. You have 60 seconds to click through following categories:

- Flats/housing
- Rooms/shared
- Sub-lets / temporary

Pay attention to the ads you see!

8.3.2 No Adaptation - Task 1.2. Finding a shared room in Espoo (test phase)

You remember you saw an ad about a shared room in Espoo from June on. You have to find it from the home page as quick as possible.

8.3.3 Reorder - Task 2.1. Check out furniture ads (learning phase)

In your new flat, you need some furniture. You have 60 seconds to click through the following categories and familiarize yourself with the ads:

- furniture
- appliances
- antiques
- general
- household

8.3.4 Reorder - Task 2.2. Finding paint

You got a call from the room renter that you will get it if you paint the walls white. You remember you saw an ad about paint in the ads - find it from the home page as quick as possible.

8.3.5 Highlight - Task 3.1. You are curious what services are offered in Helsinki area - you have 60 seconds to click through the different categories and familiarize yourself with the ads:

- Housing / Real estate
- Cars/Automotive
- Computer

8.3.6 Highlight - Task 3.2. A friend calls you that they will come to visit you from Estonia via car and wants to leave it somewhere. You remember you saw an ad about a cheap parking place in Helsinki âĂŞ find it from the home page as quick as possible.