

Identifying Different Visual Patterns in Web Users Behaviour

Martinho Gonçalves*
UTAD
INESC Porto

Tânia Rocha†
UTAD

Luís Magalhães‡
UTAD
INESC Porto

Emanuel Peres§
UTAD
CITAB

Maximino Bessa||
UTAD
INESC Porto
Warwick digital lab. Institute

Alan Chalmers||
Warwick digital lab. Institute

Abstract

On the internet one can find all kinds of information and perform a multitude of tasks such as searching for information about any subject, purchase goods and services, or enquiring about a particular company. When navigating the web, our focus (where we look) can be captured involuntarily through a visual stimulus, or may be directed in a controlled manner depending on the task that we are performing. The layouts of web sites range from simple to highly complex. When creating new layouts, knowing what information is most important when a user is performing a task and which areas are supposed to attract his/her attention are important issues in order to make a web site easier to interpret and consequently their usability. In this paper we present an eye tracking study of users' visual behaviour when surfing on websites and investigate if it is possible to determine likely visual patterns. We show that the visual patterns are different depending on the task that is being performed and not on the web site where it is being carried out.

CR Categories: H.5 [INFORMATION INTERFACES AND PRESENTATION]: 1—7

Keywords: Eye tracking, Visual attention, World Wide Web, Usability

1 Introduction

The use of the Internet and technology associated with it, has become indispensable for people around the world. On the web we can find all kinds of information and perform a large number of tasks that can go from searching for information on a particular subject, the purchase of goods and services, to searching for job positions with a company, among others.

The appearance of sites where all these tasks are performed typically have different layouts and diverge from simple to very complex. Knowing what is the most important information when a user is performing and what areas are likely to attract their attention the

most, is important when developing new designs, making them easier to interpret and more usable.

If we can distinguish between the tasks and users navigation patterns, we may be able to identify the most important areas when we are developing a new site, enabling the most important information to be placed in these locations. Furthermore, this information may be used to develop a dynamic web site interface that adapts to the tasks, in order to improve web site's usability.

The purpose of this study is to verify the existence of differences in visual patterns when surfing the Internet, depending on the tasks that are being performed.

Over the last few years, several studies have been conducted on users' visual behaviour when navigating web pages. In the next section, we present the most relevant papers for the study that we performed. In section 3, the experiment that we conducted is described, and in section 4 the principal results are presented. Finally, in section 5, some conclusions are presented and future directions described.

2 Related work

Several studies have recently been conducted on users' visual behaviour when navigating web pages. Poole and Ball[Bilal and Kirby 2002] studied the usability and interaction between humans and computer using an eye tracking device. With the eye tracker results it is possible, not only to quantify the individuals interaction with the computer, but other data can also be obtained, such as its meaning when relating the results with web sites. Poole and Ball identified variables and factors, from evidences found in other studies, that indicated not only the time that individuals take to perform certain tasks, but also the possibility of different task complexities, the difficulty in obtaining information and the structuring of web pages.

In addition, several researchers have tackled the issue of searching on Internet using search engines, as this is one of the most performed tasks. Spink et al.[Spink et al. 1999] performed a study where they analyzed users' information searching behaviour on the EXCITE search engine, through a user survey. This study identified serious problems in the design of search engines, such as lack of transparency. This means that advanced searches are more difficult to complete, losing some of their benefits compared to basic searches.

Hsieh-Yee[Hsieh-Yee 2001] conducted a study on web search behaviour, through a literature review from 1995 to 2000 that included most of the studies on children and adults. The majority of the studies that focus on children only described their interaction with the web. While in the studies that addressed the adults, the researchers' goals were mainly to describe patterns. Some studies have also investigated the effects of several factors in search behaviour, including the organization of information and its presentation, the type

*e-mail:martinhofg@gmail.com

†e-mail:ttaniarocha@sapo.pt

‡e-mail:lmagalha@utad.pt

§e-mail:eperes@utad.pt

¶e-mail:maxbessa@utad.pt

||e-mail:alan.chalmers@warwick.ac.uk

of search task, the web experience, cognitive skills and affective states.

What distinguishes the research on adults' web search behaviour is the use of multiple methods of data collection. Research on web search behaviour requires that there is a commitment to examine users in their natural environmental conditions and requires a rigorous design and data analysis. Although these studies have an increasingly rigorous objective, there is still a lack of their acceptance, as well as validation. To verify that the difference in the age influences the success in conducting searches on the Internet, Bilal and Kirby conducted a study with children and adults [Bilal and Kirby 2002]. In this study, they concluded that the search on the Internet is affected by: the users' cognitive capacity, affective state and physical state, since the user gives a different answer depending on the type of task he/she performs. The study highlighted the ability of adults to reformulate the search when faced with an impasse, the type of navigation used throughout the pages and the concentration while performing a task. The cultural experience of adults is one of the factors that most distinguishes them, resulting in a higher success rate.

In 2004 a study was conducted where Granka et al. [Granka et al. 2004] analyzed the behaviour of individuals when they perform web searches, with regard to how they interact with the sites and the results they get. Having as reference the time that the individuals take to select the desired link, this study demonstrates the complexity of the mental process involved in the choice of the link, as well as the complexity of the information presented to the users. The order, in which they appear in the search results, is related to the time that users pay "attention" to the links therein. Therefore, the search is automatically limited to the behaviour of the search engine and how it presents its results.

Lorigo et al. [Lorigo et al. 2008] also studied online searching, and presented knowledge drawn from the experiments and the difficulties encountered. Comparing different search engines, they verified that all the metrics used did not show significantly different values and thus the search in the different search engines were similar in the visual process and in the complexity of reasoning. Moreover, the individuals' preferences of the search engines used was different. The main difficulties identified were: the interpretation of the eye tracker data, analyzing the interaction between individuals and the search results, and integrating eye tracker studies with usability studies. The lack of feedback and of the task's success analysis, limits the types of studies that can be done with eye trackers. Pan et al. [Pan et al. 2004] investigated the importance of visual behaviour on web pages, analyzing the order in which pages were visited, as well as their complexity. They concluded that the human visual behaviour while viewing web pages is the result of the order in which pages are viewed, of the relationship between the site types and of the gender. The results of scanpaths (path that the eyes follow during the viewing of web pages) showed that there is the possibility of relating the visual behaviour of different individuals with the visual structure of web pages.

3 Analysis of Web Visual patterns

The shape and appearance of web sites are diverse, ranging from simple layouts to complex web applications. Knowing what information is most important when a user is performing a task and which areas are supposed to attract his/her attention are important issues when creating new designs in order to make them easier to interpret. In this paper, we present a study that verifies if users performing different tasks on the Internet follow different visual navigation strategies, resulting in different visual patterns. Therefore

we have designed an experiment using an eye tracker to find out the users' visual patterns, for which we consider the following null hypothesis:

H0: "When users perform different tasks on the Internet the visual patterns are similar".

3.1 Participants

All 83 participants were University students. Only 44 of the results were considered valid, i.e. without having any missing data. This set comprised 10 females and 34 males. Eight of the participants were over the age of 30 while the remainders were between 18 and 30 years old. 28 of the participants had normal vision and 16 had corrected to normal vision. All participants use the Internet on a daily base.

3.2 Metrics

The main metrics used in eye movement's research are: fixations and "saccades," which in turn occur between anchorages. There are also several metrics derived from these key measures, including the size of the "gaze" and the distance traveled, "scanpath". Pupils' dilation and the number of times that the subject blinks can also be studied [Salvucci and Goldberg 2000].

Using this metric, we can identify the differences in viewing patterns, because it provides a quantitative measure of the dispersion of the visual field. Moreover, this metric can be used to identify if the user is more focused on some areas than others.

3.3 Procedure

All experiments were conducted individually and in a controlled environment. The room was equipped with a computer, an eye tracker, Tobii X50, in combination with Clear View 2.0.1 software to collect eye movement data. The experiments were visualized on a 17 inches monitor with a resolution of 1024 x 768 pixels. At the end of each task, the application automatically saved all the information. The web pages were accessed using the Internet Explorer browser with a 100 Mbps internet connection. The statistical analysis software, SPSS 17.0 Statistic, was used to analyze the data provided by the eye tracking system.

To verify if the experimental setup was correct, two pilot tests involving a total of 30 participants were carried out. Problems concerning the Internet connection, types of tasks performed, experiment duration, and the participants' behaviour while performing the tasks, were considered. The results of the two pilot studies showed that the way how people behaved during the course of the experiments did not allow the eye tracker to capture their visual field. Moreover, two of the tasks had to be modified because they weren't clear enough and the participants had difficulties in completing their objectives. It was also necessary to increase the time needed to perform the tasks.

In order to perform the full experiment we defined two distinct groups. In the first group, participants performed four different tasks on four different web sites. The second group was similar to the first except that tasks 2 and 3 were modified in order to compare different tasks performed on the same web site. All these tasks were performed in well-known and fairly well-used websites, which also represented the most common types of web sites that can be found on the Internet.

The assigned tasks for the first group test were:

Task 1 - Find a recipe for a dessert using the Google search engine[Google 2009]. The task ended when the user opened the Web site.

Task 2 - Explore the site of the Espresso newspaper[Expresso 2009]. The task ended when the user closed the Web site. There were no restrictions.

Task 3 - Buy a Dan Brown book at Amazon[Amazon 2009]. The task ended when the user placed the product in the shopping cart.

Task 4 - Find information about the "HP company"[HP 2009]. The task ended when the user opened the Web page with the HP company information.

For the second group, the tasks performed at the site of the Express and Amazon were changed:

Task 5 - The same as Task 1;

Task 6 - Find an official BLOG of economics at Expresso newspaper Web site[Expresso 2009]. The task ended when the user found the Web page with the economics Blog.

Task 7 - Explore the site of Amazon[Amazon 2009]. The task ended when the user closed the Web site. There were no restrictions.

Task 8 - The same as Task 4.

Thus, the two tests were:

Test 1 - Task 1, Task 2, Task 3, Task 4;

Test 2 - Task 5, Task 6, Task 7, Task 8;

After finishing the experiment, participants answered a short survey to collect information such as age, gender, eyesight conditions, web navigation experience, amongst others.

4 Analysis and Discussion

In order to verify the integrity of the information collected using the eye tracker, the existence of outliers in the data was investigated. Based on this, it was decided to remove from Task 4 the data from subject 43 and from the Task 7 the data relating to the subject 32 since the values were significantly out of place. After this outlier identification process, we began to analyze the data to see if significant differences on the subjects' saccade rate when performing the task existed.

Because we want to examine if there are different visual patterns for each test (Test 1 or Test 2), we considered the following null hypothesis - H0: "The values of saccade rate for each of the tasks are similar".

Since for each test all participants performed four tasks, we applied the paired samples test to the results of each test, see Table 1. For Test 1 the existence of sphericity, value close to 1 (0.957), can be assumed which points to a significance level of 0.001, which is less than 0.05, rejecting the null hypothesis. This means that there are significant differences between the values of saccade rate for all the tasks performed in this group. In Test 2 the existence of sphericity can also be assumed, with a value of 1, observing a significance level of zero, below 0.05, thus also rejecting the null hypothesis. There are thus also significant differences between the values of the tasks' saccade rate on Test 2.

These results indicate that there are significant differences in the value of saccade rate between the tasks, suggesting the existence of differences patterns of navigation between the types of pages considered and that it depends on the task that is taking place. However, these tests did not point out in which tasks there are such differences. Therefore, we analyzed the tasks, carrying out a comparison two by two, to identify those which are different and those that show similarities. The null hypothesis remains the same, but the analysis is between only two tasks.

Table 1: Results of the tasks comparison using paired samples.

Test	F	Comparison	SIG. (Alfa=5%)
1	7.445	T1/T2/T3/T4	0.001
	21.209	T1/T2	0.000
	2.902	T1/T3	0.111
	5.688	T1/T4	0.036
	8.533	T2/T3	0.010
	7.927	T2/T4	0.015
	0.372	T3/T4	0.553
2	16.099	T5/T6/T7/T8	0.000
	27.564	T5/T6	0.000
	46.830	T5/T7	0.000
	17.476	T5/T8	0.001
	0.046	T6/T7	0.833
	2.925	T6/T8	0.105
	4.319	T7/T8	0.050

As can be seen in Table 1, for Test 1 the comparisons made between:

-Task 1 (search in Google) and Task 2 (Explore the Web site of Expresso);

-Task 1 (search in Google) and Task 4 (Find information about HP);

-Task 2 (Explore the web site of Expresso) and Task 3 (Buy one product at Amazon web site);

-Task 2 (Explore the web site of Expresso) and Task 4 (Find information about HP);

have a significance level less than 0.05, rejecting the null hypothesis, thereby indicating significant differences in the visual patterns when performing the tasks.

For Test 2, the comparisons made between:

-Task 5 (site search Google) and Task 6 (Find a Blog in Expresso web site);

-Task 5 (search in Google) and Task 7 (Exploring the Amazon web site);

-Task 5 (search in Google) and Task 8 (Find information about HP);

-Task 7 (Exploring the Amazon web site) and Task 8 (Find information about HP);

have a significance level less than or equal to 0.05, also rejecting the null hypothesis, thus indicating significant differences in performing the tasks.

With these results we can conclude that there are clear differences in the visual patterns in using search engines, online shopping, exploring a web site and a search of information on a web site, as far as the value of the saccade rate is concerned. Moreover, it can also be seen in Table 1, for:

-Task 1 (search in Google) and Task 3 (Buy one product at Amazon);

-Task 3 (Buy a product on Amazon) and Task 4 (Find information on HP);

-Task 6 (Find a blog at Expresso web site) and Task 7 (Exploring the Amazon web site);

-Task 6 (Find a blog at Expresso web page) and Task 8 (Find information about HP);

that with the levels of the significance obtained, we cannot reject the null hypothesis, which means that there is similarity in the mental process or reasoning required to complete these tasks.

Analyzing the results between tasks of different tests allows us to compare different tasks but performed on the same site, or the same task performed on different sites. Since the participants who performed Test 1 were different from the ones that performed Test 2, these tests are independent. The same null hypothesis was considered. Applying the test One Way ANOVA of SPSS to compare independent variables the following results were obtained:

Table 2: Results of the comparison between tasks in Test 1 and Test 2.

Tasks	F	SIG. (Alfa=5%)
T1/T6	29.660	0.000
T2/T7	0.960	0.333
T2/T6	1.293	0.263
T3/T7	4.396	0.043
T4/T6	2.055	0.162
T3/T6	5.217	0.029

As can be seen in Table 2, the comparisons made between:

- Task 1 (search in Google) and 6 (Find a Blog site of the Express);
- Task 3 (Buy one product at Amazon) and 7 (Explore the site Amazon);
- Task 3 (Buy one product at Amazon) and 6 (Find a Blog site of Expresso web site);

have a significance value less than or equal to 0.05, rejecting the null hypothesis, thereby indicating significant differences with regard to the execution of tasks.

For Task 2 (Explore the web site of Expresso) and Task 7 (Explore on Amazon web site) the null hypothesis cannot be rejected, because there are no significant differences between the values of saccade rate. This implies that the task to explore web sites has similar values, regardless of the portal where these tasks are carried out. This is consistent with the results obtained by Jansen and Pooch[Jansen and Pooch 2001].

On the other hand, for Task 3 (Buy a product on Amazon) and Task 7 (Exploring the Amazon site), although these different tasks are carried out on the same site, the significance level ($p = 0.043$) rejects the null hypothesis. There are thus significant differences between the values of saccade rate, as well as different patterns of navigations for each of the tasks. Other information on the visual patterns, produced by participants when they are browsing the web sites, is the Gaze Map.

Figure 1 shows the gaze Map of one participant. A similar behaviour was observed for the rest of participants. We can see that the gazes are mostly focused on hyperlinks at the top-left of the web page, while on the right half of the page there are no gazes. People do not undertake an orderly reading but just look at the top five hyperlinks.



Figure 1: Gaze Map from Task 1 (search for a recipe on Google).

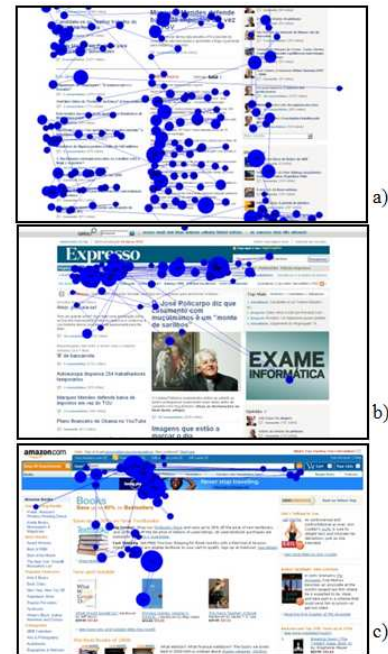


Figure 2: a) Gaze Map from [Expresso] (explore); b) Gaze Map from [Expresso] (find an Official economics Blog); c) Gaze Map from [Amazon] (Buy a Dan Brown Book).

The same observation was made for the other tasks. Differences in visual behaviour can also be verified by looking at the Gaze Maps for the different task types (Figures 1, 2) where there is a clear difference in the site navigation by users, supporting the statistical data previously presented.

5 Conclusions

The purpose of this study was to verify the existence of differences in visual patterns in the Web pages' navigation, depending on the tasks that were proposed.

In this paper we made a statistical analysis of the eye tracker data on visual patterns of web pages. In Table 1, clear differences (significance value less than 0.05) can be seen regarding the value of the saccade rate, when comparing the different tasks: searching using search engines, online shopping, exploring a site and a site search of information. Consequently, we can say that these tasks have different visual patterns.

Although in the study performed by Pan et. al [Pan et al. 2004], no significant evidence was found that visual behaviour is influenced by the type of task, in this study, as in Rayner [1998], this has been proven.

Moreover, we verified (Table 2) that the visual patterns do not depend on the site where the task is being performed.

The gaze maps for the different task types confirms the statistical results (Figure 1, Figure 2) showing that there are differences in the visual patterns when users perform those different tasks on the Internet.

6 Acknowledgements

Martinho Gonçalves has a PhD fellowship granted by FCT-Fundação para a Ciência e a Tecnologia (SFRH/BD/64013/2009).

This work is partially supported by the Portuguese government, through the National Foundation for Science and Technology - FCT (Fundação para a Ciência e Tecnologia) and the European Union (COMPETE, QREN and FEDER) through the project PTDC/EIA-EIA/108982/2008 entitled "3DWikiU - 3D Wiki for Urban Environments".

References

- AMAZON, 2009. Amazon.com Online Shopping. <http://www.amazon.com/>. [Online; accessed 23-January-2009].
- BILAL, D., AND KIRBY, J. 2002. Differences and similarities in information seeking: children and adults as web users. *Inf. Process. Manage.* 38 (September), 649–670.
- EXPRESSO, J., 2009. Expresso — Noticias da actualidade. <http://aeiou.-expresso.pt/>. [Online; accessed 23-January-2009].
- GOOGLE, 2009. Google. <http://www.google.pt/>. [Online; accessed 23-January-2009].
- GRANKA, L. A., JOACHIMS, T., AND GAY, G. 2004. Eye-tracking analysis of user behavior in www search. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, New York, NY, USA, SIGIR '04, 478–479.
- HP, 2009. Hp - Portugal. <http://welcome.hp.com/country/pt/pt/-welcome.html>. [Online; accessed 23-January-2009].
- HSIEH-YEE, I. 2001. Research on web search behavior. *Library and Information Science Research* 23, 2, 167 – 185.
- JANSEN, B. J., AND POOCH, U. 2001. A review of web searching studies and a framework for future research. *J. Am. Soc. Inf. Sci. Technol.* 52 (February), 235–246.
- LORIGO, L., HARIDASAN, M., BRYNJARSDÓTTIR, H., XIA, L., JOACHIMS, T., GAY, G., GRANKA, L., PELLACINI, F., AND PAN, B. 2008. Eye tracking and online search: Lessons learned and challenges ahead. *J. Am. Soc. Inf. Sci. Technol.* 59 (May), 1041–1052.
- PAN, B., HEMBROOKE, H. A., GAY, G. K., GRANKA, L. A., FEUSNER, M. K., AND NEWMAN, J. K. 2004. The determinants of web page viewing behavior: an eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications*, ACM, New York, NY, USA, ETRA '04, 147–154.
- POOLE, A., AND BALL, L. J. 2005. Eye tracking in human-computer interaction and usability research: Current status and future. In *Prospects, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction*. Pennsylvania: Idea Group, Inc.
- RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological bulletin* 124, 3 (Nov.), 372–422.
- SALVUCCI, D. D., AND GOLDBERG, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, ACM, New York, NY, USA, ETRA '00, 71–78.
- SPINK, A., BATEMAN, J., AND JANSEN, B. J. 1999. Searching the web: a survey of excite users. In *Internet Research: Electronic Networking Applications and Policy*, 117–128.