The evaluation of academic website using eye tracker and UEQ: a case study in a website of xyz

IOP Conf. Series: Materials Science and Engineering **703** (2019) 012049 doi:10.1088/1757-899X/703/1/012049

The evaluation of academic website using eye tracker and UEQ: a case study in a website of xyz

A H Kusumo and M Hartono
Department of Industrial Engineering, University of Surabaya, Raya Kalirungkut, Surabaya, 60293, Indonesia

E-mail: argohadi@staff.ubaya.ac.id

Abstract. The current study evaluated the academic website using eye tracking and User Experience Questionnaire (UEQ). A total of 46 participants had been recruited for this study but then the data were reduced to 33 participants due to random response errors of an item in UEQ followed by reducing in eye tracker's data. Eye tracker had a gaze plot (scan path) data and task completion time as time to first view area of interest (AOI) where these data were used to evaluate important gaze plot and whether the provision of information in website had been clearly done. The gaze plot gave insight where important link, information and picture should be put in the website. Furthermore, Kruskal Wallis and Mann Whitney U test were also applied to see the difference between each task which were used to improve sub menu's name. At the same time, UEQ data showed that only novelty should be improved to be more creative, inventive, leading edge and innovative. Thus, this recommendation provides more input in designing website for academic use.

1. Introduction

Web is playing a significant role in diverse application domains such as business, education, industry and entertainment. As a result, there are increasing concerns about the ways in which websites are developed and the quality of information delivered [1]. The growth of the web is profoundly changing the way people interact with information and with people. This has led to an expansion of opportunities for the web on different vectors, including the massive production of contents [2]. This must be taking into account in the new website development of academic website. As a case study, a website of XYZ has been evaluated. This evaluation involved the students of IE who used the website and access the important information such as lecturer data (email and area of interest), learning outcomes, downloaded forms, requirement for final project etc. Therefore, user experience measurement was required.

The first requirement for an exemplary user experience is to meet the exact needs of the customer, without fuss or bother. Next comes simplicity and elegance that produce products that are a joy to own, a joy to use [3]. Without the user experience aspect of the user, the user will leave the website that is actually important [4]. User experience measurement can be done by eye tracker data which have been collected that demonstrate that eye movements are intimately related to the moment-to-moment cognitive processing activities of readers [5]. When users reach a Web page, they can scan the page and obtain a comprehension of it in a few seconds. Cognition refers to the ability of the human mind to acquire and manage information [6] and comprises different mental processes such as attention, memory, perception, problem solving and learning [7]. By understanding sighted users' visual understanding of Web page complexity in relation with the time of task completion, important

information should reveal with reference to the cognitive effort required for interaction with that page. To use eye tracking, participants' visual attention was measured and analyzed based on the gaze plot (scan path) and heat maps eye gaze analysis [8]. Other than using UEQ, user experience can be measured using user experience questionnaire (UEQ). UEQ has 6 scales, consisting of Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation, Novelty [9]. Eye tracking's method is called objective measurement, while UEQ is referred to as subjective measurement [10]

The current study focused on evaluating XYZ new website as an example of academic website by analyzing eye tracking and UEQ. Gaze plot (scan path) of eye tracker's data will be used to analyze navigational behavior and mean difference in time to completion task as time to first view can be utilized to develop sub menu's name. At the same time, mean value of each scale in UEQ will be used to develop interface of the website. In this study, conformity of cognitive aspect between eye tracking and UEQ was also analyzed.

2. Materials and methods

This research used two methods, that is, eye tracking as an objective measurement and user experience questionnaire as a subjective measurement. Experimental procedure for eye tracker is explained below:

Eye movements were recorded by using Gaze point eye tracker. This eye tracker generates raw gaze-point location data at the camera field rate of 60 Hz. The position of GP3 was below the screen, centered and as close to the bottom edge of the screen as possible. This was approximately arm's length (65 cm) from the user. The device was approximately 40 cm below eye level and pointing a tan upwards angle towards the eyes [11]. Obtained data were analyzed using the Gaze point Analysis system. It provided powerful method for collecting and analyzing eye-gaze data. Visualization could be performed through gaze replay, gaze plots, heat maps, and area of interest (AOI). Gaze replays enabled to watch the test session video again with the user's eye motion overlaid over a recording of the changing computer screen image. In these, a blue dot could be seen or user's eye, moving around a page. Gaze plots are another very valuable analysis tool in eye tracking technology. These compile the eye gaze of one user on one page. These do not combine more than one page or user in the representation. Light blue dots on the page show a number of important data items, including: 1) where the user's fixations were, 2) numbers in the dots depicting the order in which the user looked at the items, and 3) the size of the dot denoting how long the user looked at the item. Larger dots mean longer looks. There is another analysis tool in eye tracking technology that deserves a mention here, called Areas of Interest, AOIs, or Look Zones. This feature is meant to help with doing quantitative analysis [12].

This study recruited 46 participants based on Nielsen theory which stated that the minimum participant of eye tracking usability study is 6 Qualitative eye tracking (watching gaze replays) and 20 for Quantitative user testing [12].

This study used XYZ website as a research object. Participants were asked to fulfill 5 tasks. Task 1: participants were asked to explore the website for 30 seconds. This task aimed to enable participants to assess the appearance of the website and to find out the elements on the website that attract the attention of participants through gaze plot (scan path). Task 2 was to find lecturer data, task 3 to find learning outcome, task 4 to find courses assignment's form, and task 5 to find Final Project requirement. Another 4 tasks will give a prior knowledge which area can be found faster. The matrix used was time to first view or fixation which measured the length of time taking for the respondent to reach AOI for the first time. Hence, we can evaluate whether the important information can be seen faster. Vertegaal explained [13], fixation information can be used to measure the attention that individuals have paid to stimuli.

The UEQ contains 6 scales with 26 items [9] as explained below. The questionnaire and the analysis can be downloaded online [14]. Attractiveness means overall impression of the product, perspicuity means how easy to get familiar or learn the product. Efficiency means solving the tasks without unnecessary effort, while dependability means feeling in control of the interaction. Stimulation means how exciting and motivating it is to use the product, while novelty means how innovative and creative the product is.

Mean values between -0,8 and 0,8 represent a more or less neutral evaluation of the corresponding scale. Values > 0,8 represent a positive evaluation and values < -0,8 represent a negative evaluation. The range of the scales is between -3 (horribly bad) and +3 (extremely good) [9].
To detect such more or less random or not serious answers, a simple heuristic was used. The idea to detect random or not serious answers is to check how much the best and worst evaluation of an item in a scale differs. If there is a big difference (>3), this is seen as an indicator for a problematic data pattern. Such situations can also result from random response errors or a misunderstanding of an item. Thus, it makes no sense to consider a response as problematic if this occurs just for a single scale. Answers will be removed from the data set that shows a value of 3 or higher in the Critical [9] which is then followed by eliminating the data from eye tracker.

To compare cognitive aspects between eye tracker and UEQ, the average first time to view of 4 tasks was analyzed, then it was compared with mean value of each cognitive scale in UEQ (we suggested perspicuity, efficiency, dependability as a cognitive scale). Nielsen [15] explained, If the web page survives 10 second judgment, users will look around a bit. However, they are still highly likely to leave during the subsequent 20 seconds of their visit. Only after people have stayed on a page for about 30 seconds does the curve become relatively flat. People continue to leave every second, but at a much slower rate than during the first 30 seconds. Therefore, we used 30 second as a standard time. When average time to first view was more than 30 seconds, but the UEQ result had high mean value, we called it discrepancy for cognitive aspect.

The differences in the completion time for tasks 2,3,4,5 (H1) were also analyzed. Since data were not normally distributed, we used Kruskal Wallis test which is a nonparametric alternative to a one-way ANOVA [16]. Further analysis using Mann Whitney U Test was used to see the difference between each task [17]. Statistical significance was analyzed through the use of SPSS 18.

3. Results

As was stated in the methods, a total of 46 participants were recruited for this study but the data were reduced to 33 participants due to random response errors or a misunderstanding of an item in UEQ [9] which then followed by eliminating the data from eye tracker.

The first set of analyses examined the gaze plot from eye tracker's gaze replay in response to task 1. We presented gaze plot as a navigational behavior started from path 1 as the most first viewed area in website, then continued to path 2, path 3, path 4, and path 5, respectively. As shown in Fig. 1, The most first viewed area in website sequentially were 1) home page which then the participants scrolled down through the picture which containing the news link, 2) profile as main menu, 3) each of main menu which was explored by participants.

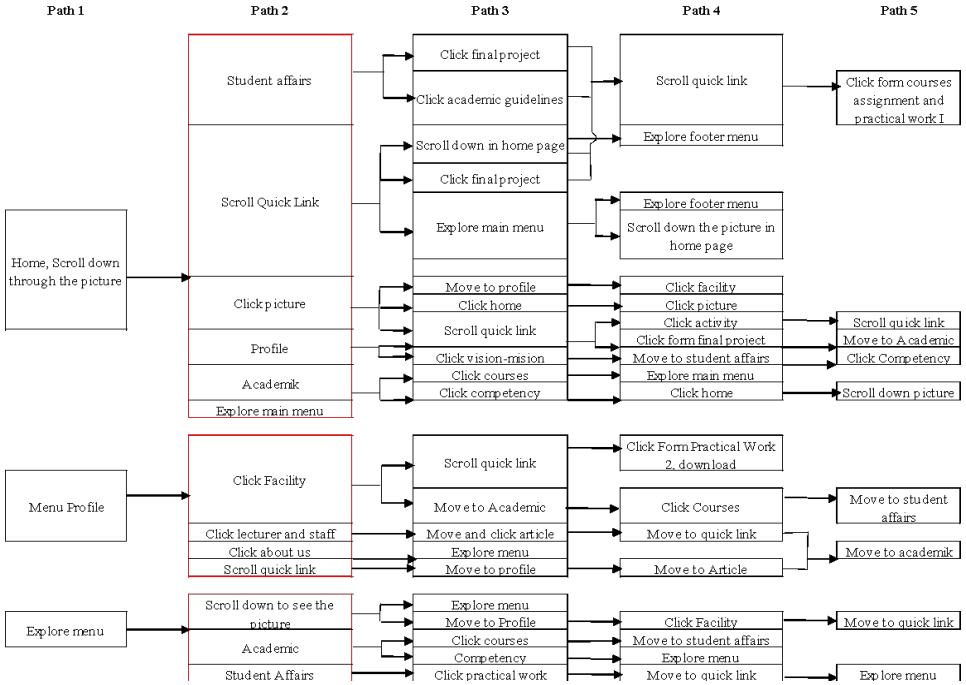

Figure 1. Gaze plot analysis of eye tracker (navigational behavior)

While scrolling on the homepage, the gaze moved to the quick link. This was mostly done not only on path 2 but also on the next path (3,4, and 5) since this quick link is located on all pages. While participants explored the quick link, they would click on the assignment and practical courses I form, the final project form, and practical work 2 from. It can be seen that quick links have an important role here. Gaze plot to profile menu could be seen in path 1, 2 and 3. In this menu, sub menu facility and lecture and staff were the most clicked and viewed. The Academic and Student Affairs were the next main menu frequently seen. After looking at the academic main menu, participants would click sub main menu, that is, courses and competency, whilst in student affairs main menu, participants clicked final project, guidelines practical work, practical work I. They also explored the menu one by one in the first path. The picture containing the news link in the home page was also clicked by some participants.

In response to UEQ as shown in table 1, perspicuity, efficiency dependability got high mean value since it was above the neutral value 0,8. At the same time, stimulation had a high value while novelty had a value that was close to a neutral value where improvement will be needed.

Table 1. UEQ result

UEQ Scales	Mean	Variance	UEQ Scales	Mean	Variance
Attractiveness	1,737	0,34	Dependability	1,833	0,35
Perspicuity	1,432	0,87	Stimulation	1,742	0,30
Efficiency	1,833	0,32	Novelty	0,811	0,67

The obtained completion time of 4 tasks from 33 participants can be seen in table 2. Completion time was obtained by putting area of interest (AOI) in the destination website page for every task that called time to first view. It can be seen that tasks 2 and 5 consumed less time than tasks 3 and 4 since task 2 and 5 have familiar sub menu's name according to participants. Sub menu for task 3 was considered unfamiliar since participants were unaware that learning outcome was a part of the competency as well as practical work 1 form was also part of the document.

Table 2. Eye tracker result in response to task 2 to 4

Task	Time to 1st View (s)	Task	Time to 1st View (s)
Task 2	16,826	Task 4	31,051
Task 3	42,897	Task 5	17,372

To see the differences between each task, Kruskal Wallis were applied. These tests revealed that there were significant differences between time for completing each task since asymp. Sig 0,00004 (asymp. Sig <0,05). To understand which time of the task was different, Mann Whitney U test was applied. The result of the test can be seen in table 3. It highlighted that time in completing task 2 and 5 had no significant difference as well as task 3 and 4.

Table 3. Mann whitney U test result for each task

Task	Mean Rank	Asymp. Sig. (2-tailed)	Conclusion
2 vs 3	24.27 vs 42,73	0,00009	Significant difference
2 vs 4	27.35 vs 39.65	0,009	Significant difference
2 vs 5	33,73 vs 33,27	0,923	No significant difference
3 vs 4	36,79 vs 30,21	0,164	No significant difference
3 vs 5	42,21 vs 24,79	0,0002	Significant difference
4 vs 5	39,97 vs 27,03	0,006	Significant difference

Moreover, verification of cognitive aspects in eye tracker and UEQ has been done. As stated in the methods, 30 seconds was determined as standard time. It was interesting to note that the lowest discrepancy was 30,30% between eye tracker's result and scale of perspicuity in UEQ. The detailed results were presented in table 4. Our experiments were in line with the meaning of perspicuity itself which showed how easy the participants get familiar or learn the website. The findings above gave recommendations of website evaluation that can be seen in table 5.

Table 4. Discrepancy between eye tracker and cognitive aspects of UEQ results

Cognitive scale in UEQ	Respondent Discrepancy	Percentage (%)
Perspicuity	10	30,30
Efficiency	18	54,54
Dependability	18	54,54

Table 5. Recommendation

Element	Recommendation
Gaze plot	● As discussed above, some participants had difficulty finding the learning outcomes (LO) and forms needed. Therefore, since LO is also important for the users, we suggest to put LO link in home page as a picture. As analyzed in path above, participants frequently scrolled down the home page through the picture. Therefore, link as a picture such as important news, accreditation, student's life, and upcoming events should be added to attract more user see this page. ● Downloaded form should be put in sub menu according to the needs. Final project form should be put in the sub menu final project while practical work form should be put in sub menu practical work. We also suggest to add more sub menu in practical work which are Practical work 1, 2, and courses assignment form. ● Quick link placement is very useful. We suggest adding more important links such as competency, curriculum, courses and eliminate less important link such as archives and categories. Practical work 1, 2, and courses assignment form will remain here.
Novelty	● Website needs to be designed in a more creative, inventive, leading edge and innovative way. ● Each page of the website should be designed more colorfully and not monotonous and more pictures should be added. ● Font should be changed from robot to more pleasant and web safe font such as arial,

AvenirNextLTPro-Regular, verdana, etc.
- Paragraph arrangement should be improved. At this time, paragraph was looked too monotonous hence it was difficult to read.

4. Discussions

The aim of this study is to evaluate academic website using eye tracking and UEQ. As analyzed in the result, the data can support this evaluation. Eye tracker as an objective measurement was used to analyze gaze plot that showed navigational behavior of the participants. This is important to see how they explored the website. A subjective measurement using UEQ showed that appropriate recommendations were needed for near neutral value of scale.

Our study failed to prove that there is a 100% conformity of cognitive aspect between eye tracking and UEQ. This can be related to standard time setting and the type of assignment given. In relation to standard time as explained by Nielsen [15], users often leave web pages in 10–20 seconds and continue to leave every second, but at a much slower rate than during the first 30 seconds. If we determine standard time was 10 - 20 seconds, level of discrepancy will be higher, more than 30,30%. Therefore, further experimental investigations are needed to estimate this standard time. In addition, further studies are also needed to explore what kind of task used for eye tracking studies which has conformity with UEQ cognitive. To avoid elimination of participants data, analysis need to be done immediately after the participants complete the UEQ.

5. Conclusions

The website is currently very important for various fields, especially education. A lot of information can be obtained by students especially from the website so that an appropriate website evaluation and development is needed. This study successfully evaluated websites using two methods, namely eye tracking as an objective measurement and UEQ as a subjective measurement. This paper has highlighted eye tracker data, that is, gaze plot and time to first view. Gaze plot could give insight where to put important information, picture or link in the website. Kruskal Wallis and Mann Whitney U test was applied to understand significant difference of time to complete each task which was used to give recommendation of changing and adding sub menu. At the same time, UEQ data showed that only novelty which had a value near neutral, hence website needs improvement to be more creative, inventive, leading edge and innovative.

In this study, cognitive level of discrepancy between eye tracker's task and perspicuity in UEQ was 30,30% due to 30 seconds time standard. This has not been able to prove 100% of the conformity of the cognitive aspect between eye tracker and UEQ. Despite the fact that there are limitations of measurement of cognitive aspects, the further studies need to develop standard time and type of task for eye tracking studies to measure cognitive aspects.

References

[1] Lopes R and Carrico L 2008 *Proc. Int. Cross Disciplinary Conf. on Web Accessibility (Beijing)* (New York: ACM) pp 5-14
[2] Iqbal M and Warraich N F 2012 *Pak. J. Libr. Inf. Sci.* **13** 1-11
[3] Norman D and Nielsen J 2006 The definition of user experience (UX) Online: https://www.nngroup.com/articles/definition-user-experience/
[4] Izabal S V, Aknuranda I and Az-Zahra H M 2018 *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* **2** 3224-32
[5] Rayner K 1998 Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* **124** 372–422
[6] Tsianos N, Germanakos P, Lekkas Z, Mourlas C and Samaras G 2009 Eye-tracking users' behavior in relation to cognitive style within an e-learning environment *Proc. IEEE Int. Conf. on Adv. Learning Tech. (Riga)* (Piscataway: IEEE) pp 329–33
[7] Solso R L, MacLin M K and MacLin O H 2004 Cognitive Psychology vol 7 (Boston: Allyn & Bacon)
[8] Nisiforou E A, Michailidou E and Laghos A 2014 *Proc. Int. UAHCI* (Crete) (Basel: Springer) pp 46–57
[9] Schrepp M 2019 User Experience Questionnaire Handbook Version 7 Online: https://www.ueq-online.org/Material/Handbook.pdf

[10] Mahardhika W, Wibirama S and Ferdiana R 2016 Analisis parallax scrolling pada web storytelling dan online shop menggunakan eye tracking dan kuesioner user experience *Electronic Theses and Dissertations Gajah Mada University*
[11] Hennessey C, Tam J and Robillard J 2014 Quick start guide Gaze point research inc Canada
[12] Pernice K and Nielsen J 2009 How to Conduct Eyetracking Studies Nielsen Norman Group Online: http://www.nngroup.com/reports/how-to-conduct-eyetracking-studies/
[13] Vertegaal R and Ding Y 2002 *Proc. Conf. on Computer Supported Cooperative Work* (New Orleans) (New York: ACM) pp 41–8.
[14] Schrepp M 2012 User experience questionnaire Online: https://www.ueq-online.org/
[15] Nielsen J 2011 How Long Do Users Stay on Web Pages? Online: https://www.nngroup.com/articles/how-long-do-users-stay-on-web-pages/
[16] Ghoodjani A 2016 Why should I use a Kruskal Wallis Test? (Tehran: Avicenna Research Institute)
[17] Hidayat A 2017 Tutorial Uji Mann Whitney U Test dengan SPSS Online: https://www.statistikian.com/2014/04/mann-whitney-u-test-dengan-spss.html/amp