Towards Automated Web Navigation and Search Skill Assessment: An Eye-tracking Study on the Skill Differences

Jakub Kubanyi, Patrik Hlavac, Jakub Simko and Maria Bielikova
Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava, Slovakia
{xkubanyi, patrik.hlavac, jakub.simko, maria.bielikova}@stuba.sk

***Abstract*—Knowledge about user's web literacy is important in automated adaptation and personalization of websites. It is also beneficial for website user experience studies. Navigation and search skills are significant constituents of the web literacy. Unfortunately, their automated assessment has so far been hard to achieve. This study investigates possibilities of a user's skill assessment based on eye-tracking data. We investigate eye-tracking metrics usable for navigation and search skill assessment. We establish a classification of users based on their skills. The paper presents an experimental study with 28 participants. Participants solved various web-search tasks, where the skill differences were manifested. We were able to detect these differences using metrics such as fixation count and time to first fixation on areas of interest.**

I. Introduction

Usability study is one of the basic tools for evaluating the design of websites. The quality of the design has major impact on the interaction of users with the website. However, there are other factors, such as the level of user skills, which influences the interaction. The relevant skill set comprises primarily the *web literacy* skills as well as domain-related knowledge as well as general computer skills. Skilled (expert) users solve tasks and overcome issues more easily while novices struggle. Users' skills influence their every-day use of Web, but also influence the outcomes of usability studies (which usually use small number of users). For example, a flawless pass through a study scenario may be a consequence of bad participant selection, even though the design of the tested page is bad. If, however, the information on user skills was available, the study outcome could be corrected.

In this paper, we investigate the possibilities of *measuring the web literacy skills*. Specifically, we focus on the impact of search and navigational skill of individual users on their interaction with websites. We strive to gain a comprehensive view on the impact of these skills. *Our core research method is the analysis of eye-tracking data*, where we seek distinctions between beginner and expert user behavior.

The web literacy is a summary of a variety of skills, each of which can be viewed separately [1]. Web literacy is a trending term[1], yet there are no standardized ways of measuring it (either through explicit questionnaires or indirectly from user behavior).

A questionnaire is a straightforward way of measuring personal characteristics. However, it also takes valuable study time. It is thus desirable to investigate ways of measuring personal characteristics (e.g. web literacy) indirectly from user behavior in other tasks. For this, the best option would be to use the same tasks for which a study is organized in a first place.

User behavior can be evaluated with the help of various sensors designed to closely monitor user behavior [2]. One of these devices is an eye-tracker – a gaze sensor. Eye-tracking has become a valuable technology for user experience researchers [3]. Modern eye-tracking technologies are able to track gaze fixations in a very detailed way. Tracking a user's behavior during his interaction with an information system can bring us many insights into the processes beyond these activities and about the user of the system as well. *The potential of eye-tracking motivated us to experiment with it in detection of user web literacy.*

With the aim of analyzing the relationship of gaze and web skills we organized a study with 28 participants. In this study, the participants were asked to perform several common web tasks found in many usability studies (e.g. information search). We recorded all user interactions with the web (including gaze). The study also included our own web-navigation skill questionnaire (there are no standard ones), results of which we used as reference information. After the study, we were able to see the differences in web navigation skills (between novices and experts) manifested in the behavior data.

II. Related work

There are several experimental studies focused on evaluation of users literacy [4], [5], [6], [7]. We are mentioning these works because just like us they studied web literacy on different types of users.

One of the experimental studies investigated literacy differences between genders [6]. Related metrics were: web

[1]It is also a fairly often mentioned term defined by several organizations, such as European Unions commission for Digital Competencies; International Society for Technology in Education, United Nations Educational, Scientific and Cultural Organizations for Global Media and Information Literacy; Next Generation Science Standards.

978-1-5386-8225-8/18/$31.00

search efficiency, various activities count, time spent in activity and even anxiety during search process. Study contained experiment focused on web search. A questionnaire related to everyday Web usage was included. After the questionnaire was filled, participants had to find answers to search tasks. Participants did not have knowledge about topic that tasks were related to. Participant's activity was recorded during the process of searching for answers. Specific types of activities were recorded along with their frequency. Every participants answer was evaluated. Relevance of these answers was judged by experts in task topics. Experts scored answers on interval from 1 to 5 points. Search Performance Index [6] was used to evaluate search efficiency.

Another experimental study investigated navigation on the Web [5]. The concept of the experiment was based on solving a series of tasks in the application: choosing the addressee, writing the message, choosing the attachment and sending the message. Navigation literacy was investigated in specific application with focus on literacy of elder users with low-literacy expectation [5]. User's study was focused on linear and hypertext navigation style. During experiment two versions of the same application have been used. Goal of the experiment was to find out in what version are users more literate. Several metrics were chosen: success rate, solution time and fixations on area of interest (AOI) visualized by heat maps. Hint section of the task was defined as AOI. The study mention that the best way to measure efficiency in literacy is success rate divided by total solution time [5]. This method is related to Search Performance Index in previous work [6].

The third experimental study provided an analysis of web search literacy [4]. The study contained an experiment with elementary school students. Along with tasks related to information search, this experiment was also focused on writing users queries. Experiment investigates mainly typed content, number of words and query modifications. We consider this experiment as interesting because it did not focus on exact location of information [4]. Instead, participants were working with search engine result page. Based on participants queries a result page was displayed. A decision if the references links were relevant or not was up to participant himself. Experiment task was to choose the most relevant reference link for the searched information. This choice had to be done in the least time as possible. The study mentions that literacy for efficient evaluation of any information relevance is based on users critical reading skill [4]. This was also mentioned in related web literacy studies [1], [8], [9], [10], [11]. Expert readers were able to reach searched information easier because they are far better in rejection of non-relevant information than less-experienced users. With less-experience users this takes longer because they are struggling with judgment if the information is relevant or not [4].

Another study aims to expand the existing research in several new areas: Instead of asking for a typical web page, the study asked about expectations for three specific web page types: online shop, a news portal, and a company web page [7]. In older study(2001), web objects had to be placed separately and unrelated to each other [7], [12]. In contrary, participants were asked to actively construct a web page as a whole in the present study. They could position the provided web objects interactively per drag-and-drop at their expected locations [7].

The present study was conducted as an online survey and as such took place in the users natural environment [7]. Finally, demographic data, experience in computer and web usage and web design expertise were assessed to build user groups in order to examine whether mental models differ between subsamples [7].

Another work presented results of a quantitative eye-tracking study with 45 participants comparing the designed visual search task to the standard conjunction search where the reaction time, number of fixations and search strategies were considered [13]. The results of the study show that searching for icons is a harder task eliciting more fixations and longer reaction times [13]. There is also study that deals with method for automatic estimation of the users interest in a webpage he visits [14].

III. Experiment design

The main content of our experiment is solving various tasks, which are focused on searching for information on the Web. Participants choose one of our 5 offered answers after the information on specific question is found. One of the answers is always correct. Other options represent distractors - the answers that can be found during the wrong searching ways. Participants can also choose options *I dont know* or *I dont remember the task question* in cases that participants give up or participants forget current task. In all our prepared tasks we are expecting minimal experience with the topic, which the task is related to.

The tasks were divided into two categories: general oriented tasks and domain oriented tasks. The former represented tasks, where participants were allowed to navigate anywhere on the web (to solve the task). The latter were tasks, where the navigation was bound to single website. We opted for this explicit distinction to cover two different cases of web navigation.

Experiment was followed by our own calibration validation. Participants were instructed to look at single point on 5 screens for 5 seconds. After all, we calculated precision for every participant.

A. Research questions

In our study, we expect differences among categories *Beginner, Intermediate, Expert* in various metrics. We expect that *time spent writing queries* and *time spent searching on opened links* are higher in the *beginner* category where are less-experienced users. Furthermore we expect that less-experienced users view and open more links than *experts* because there is a chance that they do not find desired information on the first attempts. We further investigate following questions:

- How is web-navigation literacy reflected in user behaviour, represented by set of gaze metrics?
- What are the differences in eye fixation metrics among users according to their level of navigational skill?

B. Generally oriented tasks

Generally oriented tasks are not related to specific domain. All participants start on default website https://www.google.sk/. Based on typed queries in the search bar participants open one or more links that in their opinion lead to the correct answer. In this paper we mention two examples of generally oriented tasks that has been translated to English language for the paper purpose:

1) *Find the name of the actor who has won the most academy awards during his carrier.*
2) *In what architectural style was built the Church of St. Elizabeth situated in Bratislava also known as Blue church.*

The environment allowed participants free movement around the Web, so the possibility of measuring the gaze evenly was limited in this approach. For generally oriented tasks we evaluated the following metrics: total solution time, time spent on writing queries, time spent on searching in opened links, count of correctly answered tasks, queries count, count of query modifications, viewed links count (obtained from gaze data) and opened links count.

C. Domain oriented tasks

Tasks in this category are related to specific domains. Participants start on the main pages of prepared websites in each domain. Our effort was to choose domains that are not frequently used, therefore not well known. We also chose a variety of domains (an e-shop was chosen along with book database and travelling website). In this paper we mention two examples of domain oriented tasks that has been translated to English language for the paper purpose:

1) *Find the top rated sci-fi book in book database cbdb.cz.*
2) *Find out what is the price for car parking at Bratislava airport for two days (on website zajazdy.sk).*

Domain-oriented tasks have provided more space for measurements on the same stimuli, so we used a larger range or metrics. We evaluated: total solution time, count of right answered tasks, time to first fixation on AOI, fixations before AOI, First Fixation Duration on AOI, Total Fixation Duration on AOI, Fixation Count on AOI, Total Visit Duration and Visit count on AOI. Area of interest in this experiment represents area where the searched information is located or key area that leads to this location. The search depth is relevant in cases when searched information is located on multiple paths.

D. The results interpretation approach

As well as in the related studies; we measure search efficiency through Search Performance Index[5][6].

$$SPI = (succesRate/time) \tag{1}$$

Participants were divided into three categories beginner, intermediate and expert based on their SPIs. Division is done by K-means algorithm. Each cluster computed by this algorithm represents one of the categories (beginner, intermediate, expert). Categories in our study are similar to categories presented in related web literacy studies[15][10]. Finally, we compared the calculated score from our web-navigation skill questionnaire for every single participant.

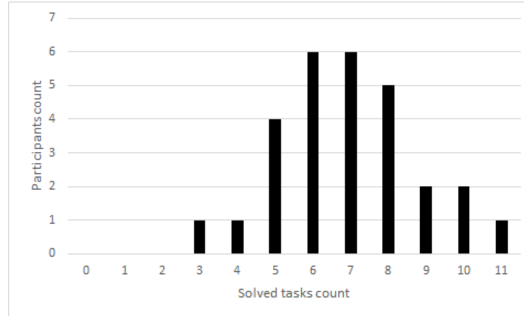

Fig. 1. Success rate of correctly answered tasks.

E. Environment

Experiment was held at the Faculty of Informatics and Information Technologies at Slovak University of Technology in Bratislava. Experiment took place in User Experience and Interaction Research Center [2] group laboratory using computers equipped with Tobii X2-60 eye-trackers. Fixations were computed from raw gaze points by Tobii I-VT Filter.

F. Participants

Participant sample was N = 28, avg. age = 17. Every participant had assigned ID for matching data from different experiment environments. After the session was over we interviewed participants about overall difficulty of our tasks. Every participant was rewarded with a small gift. After data validation, we found out that one participant was not classified by eye-tracker. This data was not included in further analysis.

IV. RESULTS

At the beginning of quantitative analysis we firstly summarized count of correctly answered tasks, both for generally and domain-oriented tasks. To sum up, only one participant successfully accomplished all of 11 tasks. On the other side, the smallest number of accomplished tasks was 3. Average number of correctly answered tasks among all participants was 7. We provide the histogram of success rate of this metric (see Figure 1).

We have also summarized general success rate of tasks. Two of them was successfully solved by 85% participants. We have marked these tasks as easy because even less-skilled users could handle them. On the other hand, there were two tasks where more than 60% participants failed. We have marked these tasks as difficult and we consider them as distinction among experts and the rest of the users. Other tasks were successfully solved by 70% participants. These are tasks of intermediate difficulty.

[2]http://uxi.sk

TABLE I. In generally oriented tasks, beginner, intermediate and expert users differed from each other on multiple metrics. Significant differences are highlighted

Metrics/Category	Beginner	Intermediate	Expert
Time spent writing queries[s]	**111,29 (SD = 39,24)**	93,64 (SD = 36,46)	**68,17 (SD = 21,57)**
Time spent searching on links[s]	111,86 (SD = 60,88)	110,69 (SD = 92,66)	86,67 (SD = 50,01)
Solution time[s]	223,14 (SD = 86,61)	232,81 (SD = 91,92)	154,83 (SD = 53,70)
Solved tasks count	**1,14 (SD = 0,69)**	**2,75 (SD = 0,89)**	**3,17 (SD = 0,75)**
SPI	**0,30 (SD = 0,14)**	**0,73 (SD = 0,12)**	**1,27 (SD = 0,19)**
Viewed links count	3,39 (SD = 2,14)	2,68 (SD = 1,36)	2,33 (SD = 1,03)

TABLE II. Experts spend consistently lesser amount of time on tasks than intermediate and novice participants, as can be seen from average total fixation duration. Significant differences are highlighted.

Task	Beginner	Intermediate	Expert
1.	6,18 (SD = 4,98)	4,46 (SM = 3,38)	1,52 (SD = 1,19)
2.	4,26 (SD = 1,90)	2,98 (SD = 2,65)	3,32 (SD = 2,90)
3.	13,12 (SD = 10,52)	**9,25 (SD = 2,75)**	**4,3 (SD = 1,31)**
4	20 (SD = 16,45)	15,37 (SD = 9,96)	8,22 (SD = 2,33)
5	5,23 (SD = 3,11)	4,45 (SD = 2,33)	2,85 (SD = 0,11)
6.	2,86 (SD = 1,35)	2,44 (SD = 1,10)	2,46 (SD = 1.10)

A. Results in generally oriented tasks

For analysis of generally oriented tasks, participants were categorized based on their SPI. We have found significant difference in *time spent on writing queries* metric among categories - beginner and expert what was also confirmed by ANOVA. Expert users spent much less time on writing queries as compared to beginner users (p-value: .035853 <.05). Experts have also tendency to spend less time by searching the links than beginners (see Table I).

B. Results in domain oriented tasks

Based on the analysis of the domain-oriented tasks, we have found differences among chosen metrics in our categories – beginner, intermediate and expert. The differences have been shown in following metrics: *total fixation duration on AOI* (also NOT on AOI as well), *fixations count on AOI* (also NOT on AOI as well) and *visit count*. Our primal expectation about differences in individual categories has been confirmed. In following paragraphs we bring a summary of metrics and differences among categories for every domain-oriented task. AOIs in this summary contain areas on websites that were crucial for certain task. Specific AOIs were resembled to key areas of a tasks, which led to task accomplishment.

In metric *total fixation duration* a gradual decrease of values across categories have been noticed (see Table II). This means that participants who have been classified as beginners had tendency to spend more time in AOIs than participants who have been classified as intermediate. Above all, participants in expert category spend the least time in these AOIs. This pattern has been noticed in 4 out of 6 tasks. We also noticed that participants classified as expert reached up to 3-4 times less values in *total fixation duration* metric as compared to participants in beginner category in two tasks. In one of our remaining tasks were values in *total fixation duration* metric similar in each category. This similarity was probably occurred because of easy difficulty of this task, which has been accomplished nearly by 80% of participants.

The metric *fixation count* is in correlation with previous metric *total fixation duration*. In *fixation count* metric have

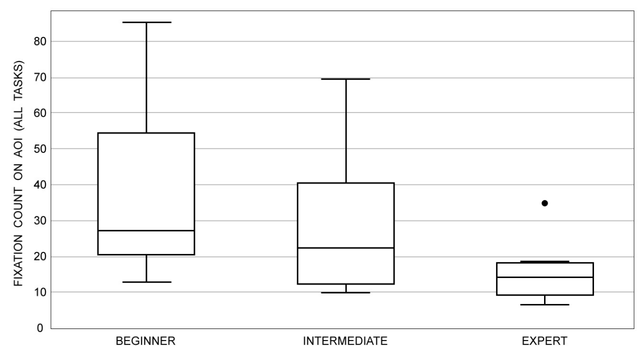

Fig. 2. Box plot: fixation counts across categories show less attention spent by experts on important parts (AOI).

TABLE III. Experts have consistently lesser interaction on tasks than intermediate and novice participants, as can be seen from listed metrics. Significant differences are highlighted

Task	Beginner	Intermediate	Expert
FIXATION COUNT			
1.	29,44 (SD = 20,60)	21,1 (SD = 17,03)	6,67 (SD = 5,50)
2.	20,29 (SD = 8,47)	12,45 (SD = 7,87)	18,33 (SD = 16,65)
3.	54,73 (SD = 44,46)	**40,6 (SD = 10,24)**	**15,33 (SD = 3,05)**
4	85,36 (SD = 68,16)	69,55 (SD = 44,08)	35 (SD = 13,00)
5	24,88 (SD = 14,56)	23,89 (SD = 10,14)	13 (SD = 1,73)
6.	12,78 (SD = 5,51)	9,91 (SD = 3,70)	9,33 (SD = 5,50)
FIXATION COUNT NOT ON AOI			
1.	**95,33 (SD = 55,68)**	113,1 (SD = 90,33)	**20,33 (SD = 14,50)**
2.	60,86 (SD = 26,72)	39,09 (SD = 30,28)	34,33 (SD = 39,70)
3.	38,36 (SD = 29,60)	25,6 (SD = 12,45)	24,67 (SD = 9,07)
4	Not measured	Not measured	Not measured
5	46,5 (SD = 30,30)	23,89 (SD = 22,99)	8,33 (SD = 8,50)
6.	69,56 (SD = 43,22)	37,27 (SD = 18,63)	26 (SD = 3,46)
FIXATION DURATION NOT ON AOI [s]			
1.	19,07 (SD = 11,43)	23,21 (SD = 18,25)	5,04 (SD = 3,53)
2.	11,33 (SD = 6,00)	7,43 (SD = 5,98)	6,53 (SD = 7,96)
3.	7,3 (SD = 6,16)	5,36 (SD = 2,69)	4,88 (SD = 1,93)
4	Not measured	Not measured	Not measured
5	**11,2 (SD = 6,42)**	**5,7 (SD = 5,41)**	**1,83 (SD = 1,87)**
6.	13,37 (SD = 8,18)	8,08 (SD = 4,56)	5,96 (SD = 1,75)
AOI VISIT COUNT			
1.	**10 (SD = 6,12)**	9,5 (SD = 6,91)	**1,33 (SD = 0,57)**
2.	7,57 (SD = 2,81)	4,55 (SD = 2,38)	6 (SD = 5,19)
3.	6 (SD = 4,12)	4,9 (SD = 1,66)	2,33 (SD = 0,57)
4	10,21 (SD = 8,35)	8,73 (SD = 4,88)	7 (SD = 2,64)
5	5,63 (SD = 3,29)	4,67 (SD = 3,08)	3 (SD = 1,00)
6.	5,44 (SD = 2,06)	4,45 (SD = 1,43)	3,67 (SD = 0,57)

appeared similar results as compared to *total fixation duration* (see Table III). Also the gradual decrease of values across categories have been noticed. This decrease has been shown in 5 out of 6 tasks. In addition, in several tasks *expert* category reached 3-4 times lesser values in *fixation count* metric as compared to participants in *beginner* category.

Other analyzed metrics were *total fixation duration not on AOI* and *fixation count not on AOI* (see Table III). Where we have noticed similar differences that were already mentioned in metrics above. Metrics *total fixation duration not on AOI* and *fixation count not on AOI* have shown us the fact that

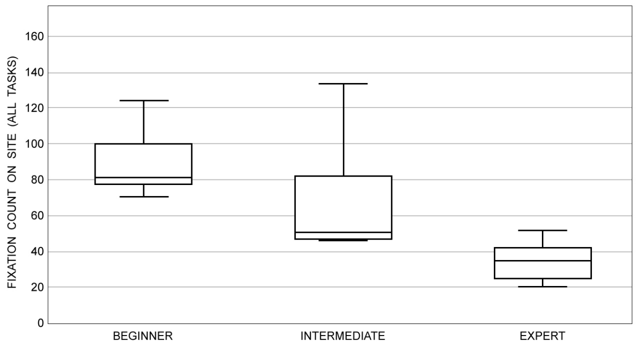

Fig. 3. Box plot: Total fixations count on all tasks.

participants classified as beginners had tendency to had more undirected search strategies and struggled with finding the key areas or searched information as compared to participants in category expert and intermediate. Values of these metrics are higher for beginners. On other hand, experts had these values lower. That may suggest that participants in the expert category have been able to identify, which key areas are important for their search and which are not - that suggest more direct search in expert category.

We have also noticed a gradual decrease of values across categories as in previous metrics. In 4. task we did not measured any values due to the nature of the task. This was very specific and it included various AOIs placed on the whole page, so the metrics NOT on AOI was pointless.

The last summarized metric is *visit count* where we find similar differences among categories (see Table III). Here we have found out that participants classified in beginner category had tendency to repetitively return to AOIs as compered to participants in other categories. On the other hand, participants classified as experts had tendency to return to these AOIs not so often. The higher tendency of repeated returns to these AOIs in category beginner may suggest that these participants struggled with understanding that these areas are crucial for their searching and they lead to searched information.

During analysis of domain-oriented tasks we have identified two types of these tasks based on the order of navigation: sequential tasks and variable sequences tasks. The sequential type of task is a task where only one way leads to the goal. Participants during solving this task are going deeper into content until they find the goal - searched information.

Based on this sinking principle the tasks are divided into phases. The first phase starts always on the main page of the site that tasks is related to. The next phases are focused on finding main key area, which leads to goal information. The last phase of the task is finding the goal information itself. We have noticed that the number of participants was decreasing gradually during the later phases. SPI values of those participants who reached the later phases were also gradually increasing (see Table IV).

We have focused on SPIs of participants in each phase of the task. In the first phase, where all participants have been included, average SPI was 0,74. Three participants did not find the book so they did not get to second phase. Number of

TABLE IV. SEQUENTIAL TASK 1 SHOWS GRADUAL INCREASE OF SPI AND LOSS OF PARTICIPANTS THAT REACHED LATER PHASES IN ORDER THEY OCCURRED (PHASE 0 REPRESENTS BEGINNING SITE OF THE TASK; PHASE 3 REPRESENTS FINAL SITE).

	Phase 0	Phase 1	Phase 2	Phase 3
Participants count	28	25	13	11
Average SPI	0,74	0,75	0,84	0,91
Web-navigation skill	68,96	69,96	68,08	68,50

TABLE V. SUMMARIZATION OF TASK WITH VARIOUS PATHS PROVIDES A NEW PERSPECTIVE ON THE ABILITY TO DIVIDE USERS INTO GROUPS

Task path	SPI	Time	Web-navigation skill
1.	0,76	28,8	69,2
2.	0,72	34,7	71,5
3.	0,71	55,6	66,3
4	0,89	53,8	72,0
5	1,26	25,0	79,0
6.	0,38	65,0	61,3

participants who found the book was 25 and they have moved to the next phase with average SPI = 0,75.

More interesting for us was the passing from phase two to phase three where 42% of participants did not pass. So in phase three we have got 13 participants. Their average SPI = 0,84. The average SPI had increased by 8% between two phases. In the final phase 7% failed to find goal information. After the last phase we ended up with 11 participants, who were able to successfully solve the task. Their average SPI = 0,91 with 6% increased between the last two phases.

Based on this summarization we classified this task as distinctive one. This means that tendency to solve this kind of task successfully was showed by participants with generally higher SPI values in domain-oriented tasks. This pattern of gradual SPI increase was noticed also in other sequential domain-oriented tasks.

We believe that we do not have to rely on basic metrics for evaluating multiple tasks. The tasks themselves can play an important role in classification of users.

The second type of tasks contain various paths that lead to solution. Participants were divided into groups according to paths which they have chosen. Every group has found searched information through different way. For example, part of participants may find searched information on navigation menu on the main page and relevant subpages during the process. Another group may find searched information in footer of the main page. There is also group of participants who find searched information by using search bar on the main page. One of the group contains participants that did not find solution. We calculated average SPIs of participants for every individual way along with the solution time and subjective skill based on questionnaire results.

V. CONCLUSION

In this work, we investigated the relationships of user web skills and their gaze behavior. We have carried out an experiment where participants performed a series of information search tasks, on which their web navigation skills were well manifested. We recorded all the behavior of the participants, including their gaze (using eye-tracking technology). As a source of reference information on user skills, we issued a web-navigation skill questionnaire. Also, we tracked the SPI (search performance index [6], [5]) of the participants. Using both the questionnaire and the SPI, we split the participants into three groups based on their navigational skills: beginner, intermediate and expert. We then sought the behavior differences between these groups over variety of metrics.

The tasks were divided into two categories: general tasks and domain-specific tasks. In general tasks we investigated user's behaviour through his search queries and interaction on the so called Wild Web. Interaction in domain-specific tasks was bound to single website.

For domain-specific tasks, we examined the differences in the several gaze metrics. We were able to confirm that behavior of novice and expert web users differ. Differences occurred in the following metrics: *total fixation duration* on the AOI (also on not on AOI), *number of fixations* on the AOI (also on not on AOI) and also *count of visits* on the AOI. A linear drop in values occurred across categories for selected metrics.

For general tasks, we inferred metric *count of viewed links* from participants gaze, however, for general tasks we focused mainly on time metrics. This type of tasks has also showed differences between participants with different skill levels. The *time spent by typing queries into a web search bar* metric showed a significant difference between the *Beginner* and *Expert* categories. This difference was confirmed by ANOVA statistical testing. Another difference occurred in the metric *time spent searching for a response on the selected line*. Beginners also tended to open and view more links than experts. In this case, the differences were not as large as for other metrics.

When analyzing the results in the domain-oriented tasks, we discovered the tasks of two types which require different user approach. The first type was a linear task type. In this type, we identified that participants with higher SPI had better progress through the tasks. The second type of tasks was a multi path task, where participants could use one of various paths to fulfill task. Our results were verified on web-navigation skill questionnaire. The results of the questionnaire show that the participants who were placed in the *Expert* category reached the highest score in our questionnaire. *Intermediate* participants reached higher score than *Beginner*.

There is still space for exploring specific issues - we would like to identify whether someone is a beginner or not (in terms of web skills). We would also like to increase our tested sample of participants as the current sample is relatively small. We could expect better results with a larger sample of participants, since some differences among the participants could be biased.

ACKNOWLEDGMENT

This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV-15-0508, the Scientific Grant Agency of the Slovak Republic, grant No. VG 1/0646/15, grant No. VG 1/0667/18, grant No. VG 1/0409/17, the Ministry of Education, Science, Research and Sport of the Slovak Republic within the Research and Development Operational Programme, ITMS 26240220084, co-funded by the ERDF, project no. 003STU-2-3/2016. The authors would like to thank for financial contribution from the STU Grant scheme for Support of Young Researchers.

REFERENCES

[1] Mozilla, "Web Literacy 2.0 online documentation," https://mozilla.github.io/content/web-lit-whitepaper, accessed: 2017-12-04.

[2] G. Slanzi, G. Pizarro, and J. D. Velásquez, "Biometric information fusion for web user navigation and preferences analysis: An overview," *Information Fusion*, vol. 38, pp. 12–21, 2017.

[3] A. Schall and J. R. Bergstrom, *Introduction to Eye Tracking, In Eye Tracking in User Experience Design*. Boston: Morgan Kaufmann Publishers, 2014.

[4] U. K. J. N. Carolin Hahnel, Frank Goldhammer, "The role of reading skills in the evaluation of online information gathered from search engine environments," *Computers in Human Behavior*, vol. 78, pp. 223–234, 2017.

[5] A. Diana Castilla and Garcia-Palacios, "Effect of Web navigation style in elderly users," *Computers in Human Behavior*, vol. 55, no. 2, pp. 909–920, 2016.

[6] M. Zhou, "Gender difference in web search perceptions and behavior: Does it vary by task performance?" *Computers & Education*, vol. 78, no. 5, pp. 174–184, 2014.

[7] S. P. Roth, P. Schmutz, S. L. Pauwels, J. A. Bargas-Avila, and K. Opwis, "Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages?" *Interacting with Computers*, vol. 22, no. 2, pp. 140 – 152, 2010.

[8] P. for 21st Century Learning (P21), "21st Century Skills Framework," accessed: 2017-12-04. [Online]. Available: http://www.p21.org/about-us/p21-framework/264

[9] N. G. S. S. (NGSS), "Science and Engineering Practices in the NGSS," 2013, accessed: 2017-12-04. [Online]. Available: https://www.nextgenscience.org/sites/default/files/Appendix%20F%20%20Science%20and%20Engineering%20Practices%20in%20the%20NGSS%20-%20FINAL%20060513.pdf

[10] UNESCO, *Global Media and Information Literacy Assessment Framework: country readiness and competencies*. UNESCO, 2013. [Online]. Available: https://books.google.sk/books?id=jivvAgAAQBAJ

[11] International Society for Technology in Education (ISTE), "Standards for Students online documentation," https://www.iste.org/standards/for-students, accessed: 2017-12-04.

[12] M. L. Bernard, "Examining user expectations for the location of common e-commerce web objects," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 47, no. 11, pp. 1356–1360, 2003.

[13] M. Dragunova, R. Moro, and M. Bielikova, "Measuring visual search ability on the web," in *Proceedings of the 22nd International Conference on Intelligent User Interfaces Companion*. ACM, 2017, pp. 97–100.

[14] M. Holub and M. Bielikova, "Estimation of user interest in visited web page," in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 1111–1112.

[15] A. Ferrar, "DIGCOMP: A Framework for Developing and Understanding Digital Competence in Europe," 2013, accessed: 2017-12-04. [Online]. Available: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC83167/lb-na-26035-enn.pdf