Assessing the potential use of eye-tracking triangulation for evaluating the usability of an online diabetes exercise system

Schaarup, Clara; Hartvigsen, Gunnar; Larsen, Lars Bo; Tan, Zheng-Hua; Årsand, Eirik; Hejlesen, Ole

Published in:
MEDINFO 2015: eHealth-enabled Health

DOI (link to publication from Publisher):
10.3233/978-1-61499-564-7-84

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

MEDINFO 2015: eHealth-enabled Health
I.N. Sarkar et al. (Eds.)

doi:10.3233/978-1-61499-564-7-84

Assessing the Potential Use of Eye-Tracking Triangulation for Evaluating the Usability of an Online Diabetes Exercise System

Clara Schaarup[a], Gunnar Hartvigsen[b,c], Lars Bo Larsen[d], Zheng-Hua Tan[d], Eirik Årsand[c], Ole Kristian Hejlesen[a,b]

[a] *Department of Health Science and Technology, Aalborg University, Denmark*
[b] *Department of Computer Science, University of Tromsø – The Arctic University of Norway, Norway*
[c] *Norwegian Centre for Integrated Care and Telemedicine, University Hospital of North Norway, Tromsø*
[d] *Department of Electronic Systems, Aalborg University, Denmark*

Abstract

The Online Diabetes Exercise System was developed to motivate people with Type 2 diabetes to do a 25 minutes low-volume high-intensity interval training program. In a previous multi-method evaluation of the system, several usability issues were identified and corrected. Despite the thorough testing, it was unclear whether all usability problems had been identified using the multi-method evaluation. Our hypothesis was that adding the eye-tracking triangulation to the multi-method evaluation would increase the accuracy and completeness when testing the usability of the system. The study design was an Eye-tracking Triangulation; conventional eye-tracking with predefined tasks followed by The Post-Experience Eye-Tracked Protocol (PEEP). Six Areas of Interests were the basis for the PEEP-session. The eye-tracking triangulation gave objective and subjective results, which are believed to be highly relevant for designing, implementing, evaluating and optimizing systems in the field of health informatics. Future work should include testing the method on a larger and more representative group of users and apply the method on different system types.

Keywords:

Eye Tracking; Usability; Diabetes Mellitus; Post-Experienced Eye-Tracked Protocol; Consumer Health Information.

Introduction

Consumer health information technology (CHIT) is a new term in the healthcare sector, which refers to a computer-based system that facilitates access to information and behaviour changes that promote health and well-being [1]. Recently, a systematic review by Jenni Cornelly et al. examine the effectiveness of technology to promote physical activity in people with Type 2 diabetes [2]. Their findings indicate that technology-based interventions to promote physical activity are effective. Nevertheless, they demonstrate a need for evidence of the sustainability of the technology [2]. In a systematic review by Yu et al. they also identify the needs of sustainability, effectiveness, usefulness, and usability regarding CHIT and management of diabetes [3].

An example of a CHIT-system, which has been tested for usefulness and usability, is the web-browser based patient health IT system named "The Online Diabetes Exercise System". The system was developed to motivate people with Type 2 diabetes to do a 25 minutes low-volume high-intensity interval training program [4]. The development of this CHIT-system was motivated by the fact that in 2014, 387 million people worldwide had diabetes and that the number is expected to rise to more than 592 million by 2035 [5].

In a previous multi-method evaluation of the CHIT-system, several usability issues were identified and corrected [4]. Despite the thorough testing, it is unclear whether all usability problems were identified using the multi-method evaluation.

To assess the amount of usability problems that were not identified in the multi-method evaluation, we have applied an extended version of eye-tracking, named Eye-tracking Triangulation, to the CHIT-system. Eye-tracking triangulation is made of two successive sessions. The first session of conventional eye-tracking with predefined tasks followed by the second session: The Post-Experience Eye-Tracked Protocol (PEEP) [6]. Eye-tracking is a technique where eye movements are recorded while users look at a stimulus [7]. In PEEP, the users have to explain their decisions and thoughts while a retrospective replay of their eye-tracking data is shown [6]. Even though eye-tracking has been used in psychology for decades, to our knowledge, eye-tracking triangulation has not been applied to health informatics systems.

Our hypothesis is that adding the eye-tracking triangulation to the multi-method evaluation increases the accuracy and completeness when testing the usability of CHIT-systems. The objective of the present study was to assess the potential use of eye-tracking triangulation for evaluating the usability of a CHIT-system, The Online Diabetes Exercise System, which, before the present evaluation, had been optimized using multi-method evaluation.

Materials and Methods

The Patient Health IT System

As already outlined, the eye-tracking triangulation was applied to the Online Diabetes Exercise System (Figure 1). The system is a web-browser based prototype and was designed, implemented, evaluated and optimized using the iterative, systematic, and holistic multi-method evaluation, which consisted of interviews, paper prototyping, heuristic evaluation and tests with users [4]. The system contains two major functions: a 25 minute low-volume high-intensity interval training

program (HIT program), which the users in this study only looked at, but did not perform, and a glucose diary.

Figure 1 - The Menu of the Online Diabetes Exercise System

Recruitment

Eight users were recruited for the study - five women and three men. The background of the participants were diverse, ranging from having short to long educations: a biomedical laboratory technician with a master in health informatics, an occupational therapist, a global IT-manager, a nurse, a boilermaker, a shipper, an MA in Danish and Psychology, and a veterinary nurse. Their age ranged from 34 to 69 years. Six out of the eight had a Body Mass Index (BMI) of more than 25 and were per definition classified as being overweight. All the users exercised daily, they were native Danish speaking people, and they reported to use the Internet on a daily basis. Five out of the eight users wear glasses, which is not a problem while tracking their eye movements.

Equipment

The users' eye movements were tracked using a Tobii Eye-Tracker X120. The eye-tracker was placed below a Samsung Monitor 21.5 inch monitor used by the test participants, as shown in Figure 2. All eye movements were recorded using the Tobii Studio software and stored together with a screen video recording, voice input, and mouse clicks for subsequent analysis [8]. Eye-movements consist of fixations and saccades. A fixation is defined as a moment where the eyes are almost motionless and it generally has a duration of 100 ms to 500 ms. A saccade is defined as a quick movement between different fixations and on average it lasts for about 250 ms, for example while reading [6]. In this study, we were only interested in fixations periods. While the test participants sat in front of the eye-tracker, their eye movements were shown simultaneously on a separate monitor, allowing the facilitator to follow the experiment and ensuring the equipment was capturing the data at all times (to the left in Figure 2). Prior to every test session, the equipment was calibrated to each participant to ensure optimal accuracy (about 0.5 cm on a monitor placed 70 cm in front of the test participant).

Study Design and Procedure

Session 1: Eye-Tracking

Initially, each user received an introduction to the purpose of the study. They then received instructions on how to seat themselves in a comfortable position in front of the Tobii Monitor, while completing the following tasks:

1. Set up an account in the Online Diabetes Exercise System
2. Search for help due to forgotten login information
3. Login to the Online Diabetes Exercise System
4. Acquire knowledge about the Online Diabetes Exercise System through an available user guide
5. Start and finish the low volume HIT program (not performing the actual training)
6. Register and review information in the glucose diary
7. Log out from the Online Diabetes Exercise System

After assenting to the instructions and the tasks, the facilitator calibrated the Tobii eye-tracker and the users began completing the tasks.

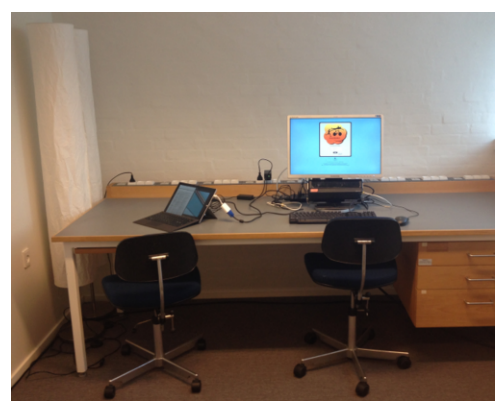

Figure 2 - The eye-tracking set up

Session 2: Post-Experience Eye-Tracking Protocol (PEEP)

A dicta-phone was turned on to collect the users' answers. The facilitator had predefined six Areas of Interests (AOIs) which would be the basis for the PEEP-session. The AOIs were: (1) *Information Text;* (2) *Forgot login/Create an Account;* (3) *Login;* (4) *Subheadings;* (5) *Colour Box;* (6) *Progress Bar.* The AOIs were spread over several different screens. The users were asked to provide retrospective protocols, meaning that while they were shown a replay of their eye-tracking data they were asked to explain their thoughts and decisions to each AOI [6]. The following details were specifically queried: long fixations; text scanning rather than reading, and failing to look at specific elements that were considered important for the task. At the end of the session, the users were asked if they had any further comments, and if not, Session 2 was complete. Session 1 took approximately 10 minutes and Session 2 took 15-20 minutes.

Data Analysis

Statistical data about the six AOIs were extracted from the eye tracking software regarding the two metrics shown in Table 1. Usability issues were identified from the statistical data. Valuable and complementary inputs to analyse the statistical data were provided by the PEEP-technique. The PEEP inputs were afterwards compared with the identified usability issues.

Table 1 – Eye-Tracking Metrics

Metric	Explanation
Time to First Fixation	The time in seconds from when the stimulus was shown until the start of the first fixation within an AOI.
Visit Count	Number of times the user fixate on an AOI.

Results

The five heatmaps shown in Figures 3 to 7 illustrate the visual behaviour of users during the eye-tracking sessions. The colour coding illustrates the total time spent looking at various areas. Green equals a shorter time, yellow medium time, and red longer time. Dashed rectangles indicate AOIs.

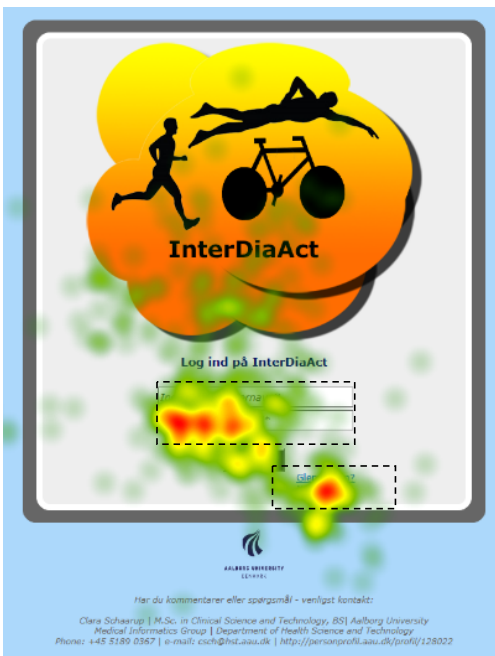

Figure 3 – The welcome screen has two AOIs: the 'Login' area and the area for 'Forgot login/Create an Account'

Figure 3 shows how the users fixated at the AOI-*Login* and the AOI-*Forgot login/Create an Account*. On average for the Login AOI, the users had 12.00 visit counts and spent 1.38 seconds before their first fixation, which can be seen in Table 2. During the PEEP session, some of the users reported that the font used was too small to identify the button 'Create an account'. Later in the session, they said that they just did not look carefully enough.

Table 2 – The six predefined areas of interest correspond to the two metrics: Visit Count & Time to First Fixation

Area of Interest; (AOI)	Visit Count; Mean(count)	Time to First Fixation; Mean(.seconds)
Information Text	10.25	17.80
Forgot login/Create an Account	5.17	10.18
Login	12.00	1.38
Subheadings	12.75	2.96
Colour Box	7.20	0.25
Progress Bar	1.00	1.33

The heatmap in Figure 4 illustrates how the users read the information about the newest research regarding low-volume high-intensity interval training. As can be seen in Table 2, the users on average had 10.25 visit counts on this AOI, the *'Information Text'*, and spent 17.80 seconds before they had their first fixation on it. In the PEEP session, the users reported that it was overwhelming with all the information on this AOI and was hard to read it all. In addition, they reported that the font size was too small. Nevertheless, some of the users mentioned that it was interesting to read. A later test of the text showed a readability index of 49, which corresponds to very difficult readings such as academic papers [9].

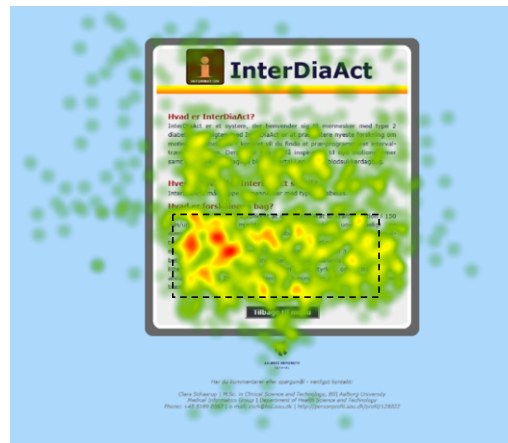

Figure 4 –AOI with the newest research regarding low-volume high-intensity interval training

Figure 5 shows the heatmap for the *Progress Bar*. On average, the users had 1 visit count and spent 1.33 seconds to identify the AOI. In the PEEP session, the users reported that the progress bar was too small and unclear.

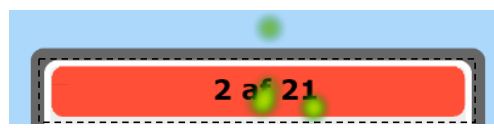

Figure 5 – While the low-volume high-intensity interval training program was running, the users were presented to a red progress-bar placed in the top of the of the screen

Figure 6 shows the heatmap for the page where users can obtain information on how to add another training program or

switch to a different type of exercise, for example from biking to swimming. The page is divided in subheadings and one of these is defined as an AOI. On average, the users had 12.75 visit counts and 2.96 seconds to first fixation on the AOI *Subheadings*. During the PEEP session, the users reported that the subheadings helped them to get an overview of the information.

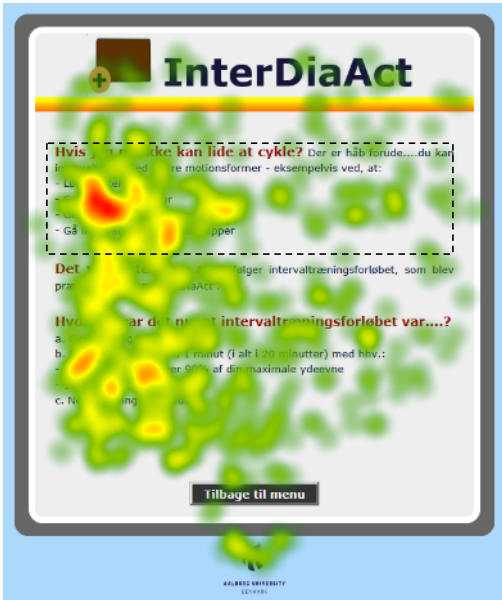

Figure 6 - The heatmap of the page where users can get information on how to add another training program or switch to a different type of exercise

Figure 7 shows the information explaining the colour scheme used during the training program – green for initial warm-up and final cool-down period, yellow for pauses between bursts of high intensity training, and red for the short periods with high intensity activity. The illustration of this at the bottom of the figure has been chosen as an AOI, the *Colour Box*. On average, the users had 7.2 visit counts and spent 0.25 seconds to first fixation at the AOI. In the PEEP session, the users reported that they preferred figures rather than text.

Figure 7 – The heatmap of the page explaining the colour scheme used during the training program

Discussion

The aim of the study was to assess the potential use of eye-tracking triangulation to evaluate the usability of the Online Diabetes Exercise System. The eye-tracking triangulation was made up of two successive sessions: The first session of conventional eye-tracking with predefined tasks followed by the second session, The Post-Experience Eye-Tracked Protocol (PEEP), the latter being based on the eye-tracking heatmaps and the six selected Areas of Interest [6].

Eye-tracking triangulation is a technique, which has been used for more than 20 years in usability studies for marketing purposes, and which is well known in cognitive psychology [7]. The technique is an attractive and powerful tool because it delivers objective results such as heatmaps and subjective results such as statements and reports that clarify potential misinterpretations of the heatmaps. The combination of objective and subjective results can be argued to make the method more valid and reliable.

In our study we saw the strength of the eye-tracking triangulation on several occasions, where the method detected usability issues that had not been found with the previously applied multi-method evaluation consisting of interviews, paper prototyping, heuristic evaluation and tests with users. For instance, we had an assumption that the participants would use the progress bar while they were looking at the low-volume high-intensity interval training program (Figure 5). However, the results showed that the users on average only had one visit count at this AOI and thus basically did not use it (Table 2). In the following PEEP session, we received an expanded explanation for this. This is an example of how eye-tracking triangulation, in our study, provided relevant information that was not provided by any of the four tests and evaluations in the multi-method evaluation.

Another example, which illustrates the strength of the triangulation, is the heatmap in Figure 4, with an AOI describing the newest research regarding low-volume high-intensity interval training. From the heatmap, we saw that the users read the information text, i.e., looked at the AOI labelled *Information Text*, but also that they looked at areas far away from the AOI several times. The heatmap did not tell us why, but the users explained this during the following PEEP session and thereby gave important information on a very relevant problem – that it was overwhelming with all the information on this AOI and that it was hard to read all. Again, this problem had not been detected by any of the four tests and evaluations in the multi-method evaluation.

The two examples illustrate how the combination of the eye-tracking and the post-experienced eye-tracked protocol can increase the accuracy and completeness of findings of usability issues in a consumer health information technology system.

From the literature, we know that it may be difficult to evaluate systems in the field of health informatics [10]. Bürkle et al. have in their study tried to assess advantages and disadvantages of different study designs, which may be potential methods to evaluate health informatics. They looked at the following designs: a randomized controlled study, a controlled study, a non-controlled study, and a simulation study and they found that evaluating health informatics depends on: the goal of the evaluation, available resources, human factors, and what type of technology that has to be examined. They concluded that a mixed approach that combines measurement of several indicators is the optimal way to evaluate health infor-

matics [10]. This conclusion is supported by the findings in the present study, which, in addition, indicates a potential added value of including eye-tracking triangulation in a mixed approach to evaluate health informatics.

Despite the promising results in our study, there were some limitations in the design. Firstly, we only used eight participants and they were probably not representative of all potential users of such a system. Secondly, the eye-tracking triangulation was only applied to one system, the Online Diabetes Exercise System, and the results may be different when applying the method to other types of consumer health information technology.

Conclusion

In conclusion, several problems, not identified previously, were found using eye-tracking – adding eye-tracking triangulation to the multi-method evaluation was found to increase the accuracy and completeness. Eye-tracking triangulation gave objective and subjective results, which are believed to be very relevant for designing, implementing, evaluating and optimizing systems in the field of health informatics.

Even though eye-tracking has been used in other areas for decades, to our knowledge, eye-tracking triangulation has not been applied to health informatics systems. Our findings suggest a future role for eye-tracking triangulation when evaluating systems in the field of health informatics.

Future work should include more studies of the potential benefits from using eye-tracking. The method should be assessed using larger and more representative groups of users and the method should be applied to a variety of different types of systems.

Acknowledgements

The authors would like to thank the users who participated in the eye-tracking triangulation sessions. The work was supported by the 7th framework EU project eWall.

References

[1] Lai YH. A Meta-Analysis on the Information Technology Application in Patients' Knowledge and Anxiety in Taiwan. Asian Journal of Applied Science. 2014;02(02):138-144.

[2] Conelly J, Kirk A, Masthoff J, MacRury S. Systematic Review or Meta-analysis. The use of technology to promote physical activity in Type 2 diabetes management a systematic review. Diabetic Medicine. 2013; 30. 1420-1432.

[3] Yu CH, Bahniwal R, Laupacis A. Systematic review and evaluation of web-accessible tools for management of diabetes and related cardiovascular risk factors by patients and healthcare providers. J Am Med Inform Assoc 2012;19:514-522.

[4] Schaarup C, Hejlesen OK. A Multi-Method Pilot Evaluation of An Online Diabetes Exercise System. MIE 2015 Proceedings, accepted, pp.5.

[5] International Diabetes Federation. IDF Diabetes Atlas Sixth Edition. International Diabetes Federation, 2013.

[6] Ball LJ, Eger N, Stevens R, Dodd J. Applying the PEEP method in usability testing. Interfaces, 67 (summer 2006), 15-19.

[7] Lai ML, Tsai MJ, Yang FY, Hsu CH, Liu TC, Lee SWY, Lee MH, Chiou GL, Liang JC, Tsai CC. A review of using eye-tracking technology in exploring learning from 2000 to 2012. Educational Research Review, Elsevier, 2013:10: 90-115.

[8] Tobii. User Manual – Tobii Studio, Version 3.2. Manual Version 3.2 Rev A. 11/2012. Tobii Technology AB.

[9] Anderson J. Lix and Rix: Variations on a Little-known Readability Index. Journal of Reading, 1983;26:6, 490-496.

[10] Bürkle T, Ammenwerth E, Prokosch HU, Dudeck J. Evaluation of clinical information systems. What can be evaluated and what cannot? Journal of Evaluation in Clinical Practice, 2001; 7:4, 373-385.

[11] Friedman CP, Wyatt J. Evaluation Methods in Biomedical Informatics. Springer. 2006.

[12] Brender J. Handbook of Evaluation Methods for Health Informatics. Elsevier Academic Press: Burlington, MA. 2006.

[13] Guideline for good evaluation practice in health informatics (GEH-HI). https://www.ghdonline.org/uploads/Eval_2011_NykanenEtAl_GuidelineGoodEvalPracHlthInfo_GEP-HI_IMJI.pdf

Address for correspondence

Clara Schaarup, Medical Informatics Group, Department of Health Science and Technology, Aalborg University, Frederik Bajers Vej 7 C2-221, DK-9220 Aalborg, Denmark. E-mail: csch@hst.aau.dk