Combining Physiological with Cognitive Measures of Online Users to Evaluate a Physicians' Review Website

Andreas Gregoriades and Olga Vozniuk
Department of Computer Science and Engineering, European University Cyprus, Nicosia, Cyprus

Keywords: Web Usability, Eye Tracking, Design Science.

Abstract: Patients' opinions is considered a new and promising way for evaluating the performance of medical practitioners. This paper presents a design science approach to the evaluation of a doctors' review website. Due to the importance of the decisions that patients might make using such websites, usability is considered as one of the critical parameter to their success. This paper reports on results from an experimental evaluation of a doctors' review website, using a combination of evaluation methods such as: eye tracking, think aloud and surveys. Results from the evaluation process highlighted a number of issues related to the information architecture which have been address during the redesign of the website.

1 INTRODUCTION

The spread of Health 2.0 (Berben et al., 2010) technologies in the last decade has made the Internet a popular place to learn and discuss health matters. As per a recent survey (Fox, 2011), 80% of users searched for health information online, and out of these, 6% have contributed to health related discussions. Online users apply crowdsourced data from online reviews to make decisions on products and services. To that end, it should be no surprise when patients use past experiences of others to decide regarding their doctor. However, choosing a competent and experienced physician is a more important decision than what merchandise to purchase online. Therefore, websites that provide such support should be designed and evaluated against different criteria not just solidly on usability.

Past research reported low usage for online review websites for physicians. However, a recent report (Galizzi et al., 2012) revealed that between 2005 and 2010 there was an increase in the number of physicians that have been rated online. A study by Davis et al. (2014), reported that 15% of individuals are aware of physician rating sites, however only 3% of them had used them. Another study reports that 32% of online users are aware and 25% usage online doctors review websites. Davis et al. (2014) also reveals that the highest rates of awareness (65%) are reported in the US population with usage (23%).

The sheer number of available doctors per specialty makes the decision of which one to consult difficult. The decision can be based on a number of criteria which in most cases are influenced by other patients. The work presented herein concentrates on the design of a non-profit website to provide patients with evaluations of medical practitioners in Cyprus. This website aims to assist patients in making a more informed decision regarding the selection of a healthcare provider. The main issues in this endeavour are: the design of a trustworthy site, the usability of the site and finally its usefulness. To be able to address these problems a framework that will define the assessment criteria and the procedure in designing and evaluating such a website needs to be specified. The underlying research framework used in this work is design science (Chatterjee 2010). The paper focuses on the evaluation phase of design science that investigates the effectiveness of an artifact and guides its re-design through changes in specification. This work is performed in accordance with the following research objectives:

1. Develop a theoretical framework and approach to evaluating websites for medical decision making.
2. Evaluate the usability, effectiveness, interface quality and information architecture of the "whattheythink.info" (WTT) website, using a combination of physiological and cognitive measures of user interaction

Gregoriades A. and Vozniuk O..
Combining Physiological with Cognitive Measures of Online Users to Evaluate a Physicians' Review Website.
DOI: 10.5220/0005518701520159
In *Proceedings of the 12th International Conference on e-Business* (ICE-B-2015), pages 152-159
ISBN: 978-989-758-113-7

3. Determine if WTT can be used to produce the outputs required for effective patient decision support.

The paper is organised as follows: firstly work related to the problem addresses is described. This is followed by the role of crowdsourcing in decision making. Next the design of the website is illustrated. This is followed with the research methodology. Next, the design of the experiment for usability evaluation is presented. Results from the experiment are presented next. Finally the paper concludes with a short discussion and future directions.

2 RELATED WORD

The main body of research in medical decision making concentrates on physicians' decision making or patients' decisions with regards to therapeutic options. On the other hand, relatively little research has addressed one of the first decisions that patients must make, which is the choice of a doctor.

Decision making is a process of making a choice from a number of alternatives to achieve a desired result (Eisenfuhr, 2011). In the consumer sector, web designers seek to understand the consumer decision process and accordingly design their websites so as to best support consumers. In this process the critical information requirements of consumers needs to be taken into consideration. The doctor selection problem is similar to the consumer problem, however the criteria are different and the seriousness of the effect of an incorrect decision are more dramatic. The steps in this process are: need recognition, information gathering, alternative evaluation, selection, post selection evaluation. Need recognition is the stage where the patient realizes the difference between desired situation and the current situation and this serves as a trigger for the doctor selection process. The next stage is search for data relevant to the decision, both from internal sources (one's memory) and/or external sources. Next comes the evaluation of alternatives that can serve the consumers' needs. Doctors review websites come to serve this stage of the decision making process. They provide patients with valuable information regarding the performance of medical practitioners and in this way support the selection stage of patients.

Crowdsourcing is one of the best techniques for gathering information from large number of people and hence is becoming a popular mean for supporting the information search stage of a decision making process. Howe (2006), defines crowdsourcing as the: "act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call". Applications of crowdsourcing have been reported useful in different fields, especially in medicine. Patients use crowdsourcing mostly online to finding solutions to health issues. At the same time doctors also use the web to exchange medical knowledge, Riley (2012). Such networks include websites like: Doximity and Sermo were doctors from different countries share clinical information (Riley, 2012). Additionally, crowdsourcing is now becoming widely used in review websites to gather information regarding online users' opinions. The link between decision-making and user generated content from crowdsourcing has been established in the literature (Brezillon, 2014). Specifically, research concludes that online reviews provide useful information to help decision makers make a more informed decision. The website presented herein uses crowdsourcing techniques to collecting information from large groups of people. To obtain patients reviews about doctors in Cyprus.

2.1 Usability as a Component of a Websites' Evaluation

Usability constitutes an important parameter in determining an artifact's success. In the context of the WTT, usability affects the decision making process. It impact on the user's ability to complete tasks from which the system is designed to support (Gabbard et al., 1999) i.e. decision support. According to Davis et al. (2003) usability is neither the only nor the most important predictor of an artifact's acceptance and usage. Usability as defined by Maxwell (2002) includes attributes such as ease of use, ease of learning, error prevention, error recovery, efficiency of performance. These constitute part of the parameter-set included in the WTT usability evaluation. Nielsen (1996) defines usability as a measure of the quality of the user experience when interacting with a web based or traditional software application. The five attributes that contribute to usability include: learnability, efficiency, memorability, errors and satisfaction. However, satisfaction cannot be limited to 'pleasant to use' and referring to Morand (2002) it has been extended to include satisfaction with the process and satisfaction with the outcomes. In our case the outcome is the selection of a suitable doctor.

3 DESIGN SCIENCE APPROACH TO THE DEVELOPMENT OF WTT

The design and development of new artifacts such as the WTT website, described herein requires a systematic approach towards artifact design, development and evaluation. This aims to assure that the artifact contributes towards resolving a particular problem. This process constitutes design science which synthesises the sciences of the artificial, engineering design, information systems development, system development as a research methodology, and executive information system design theory for the building and evaluating of IT artifacts for specific problems (Chatterjee, 2010). The work presented herein offers a framework and approach for guiding the evaluation of a special type of websites namely, doctors' review websites through an experimental case. The first part of the work focuses on the on the role of domain knowledge for defining the artifact's evaluation criteria. This is illustrated through a WTT domain analysis for the identification of critical parameters that should guide the evaluation process and subsequently the WTT redesign. The second part of the paper addresses the design and evaluation of the website. The final phase addresses the redesign of the website based on the issues that emerged from the evaluation.

The artifact in this framework is a doctors' review website. The WTT website was implemented using the joomla content management systems. It contains information regarding all medical practitioners in Cyprus, categorized by specializations. Users can search for doctors based on a number of criteria and share their healthcare experience by providing an evaluation of each doctor. To be able to evaluate a doctor, a user must be registered. All patient's reviews and ratings are anonymous. Doctors' details are modelled by joomla articles. These are augmented with details regarding doctors' reviews, medical history, specialisations, place of practise, and contact details. Each doctor is registered with the Cyprus Doctors Association. Navigation through the website is achieved using a menu bar at the top of the site. Each doctor is evaluated against a number of criteria. These were identified based on preliminary literature review on patients' opinions regarding doctors' qualities that they are seeking. The evaluation of the doctor is achieved only by registered users to enable validity of patients' feedback. Registration to the system is achieved either through form filling or social network account details. Doctors are categorized by city and specialisation. Searching for doctors is achieved using a doctors' specialization, name of doctor and city of practise. The site enables users to perform two activities, either to provide their feedback regarding a doctor once they visited a doctor or consult what other say about a doctor to make a decision about the doctor to visit. Ratings of doctors is calculated based on the score they achieved by users for all of the aforementioned criteria. Each criterion is assigned a different weighting factor. The visualisation method for the rating of each criterion is based on the popular star rating method. The website provides users the option of identifying the best doctor in different specialisation depending on patients review comments and ratings. Additional information regarding pharmacies on duty and relevant medical news is provided to users depending on their search criteria. The information requirements for the design of the user interface were based on a combination of methods, namely domain analysis and user analysis. For the former a study was conducted to identify the decision making process of patients and the latter included a series of interviews with a small group of patients in Cyprus. The selection of the users was based from a diversity of users' types in order to identify needs from different types of users that have a common problem. The study was conducted using an open interview method. During the interview the researchers were seeking to identify the information needs of patients in Cyprus. In addition the interview process helped to identify an information gab regarding the level of healthcare service in Cyprus. Specifically, the analysis revealed that patients in Cyprus base their decision, regarding which doctor to consult, solidly on reputation and word of mouth. The knowledge regarding the reputation of each doctor is mainly established through the experience of each patient with a doctor. This finding laid the foundation for the specification of the problem statement which drove the requirements elicitation phase for the identification of the system requirements for a website to be developed. The paper reports on a prototype design of the system that aimed to address these user needs. The prototype version helped to validate these requirements and accordingly guide the redesign of the website to optimise its support to those needs. Among the functional requirements that needs validation were also non-functional requirements (NFR). The NFR that we address in this study is the usability of the system. This is evaluated using traditional usability testing techniques, in combination with physiological metrics such as eye fixations and mental model

analysis using eye tracking and contextual task analysis - think aloud techniques (Figure 1). The scenarios used during the evaluation of the site were specified based on the initial task diagram and the use-cases specified at the requirements specification phase. Potential users of the website are specified as people who have a medical condition and are seeking advice prior to deciding on the selection of the doctor to consult. The theoretical assumption of patient decision making is different from the decision making in the retail sector and this was taken into consideration during the specification of the requirements for the system.

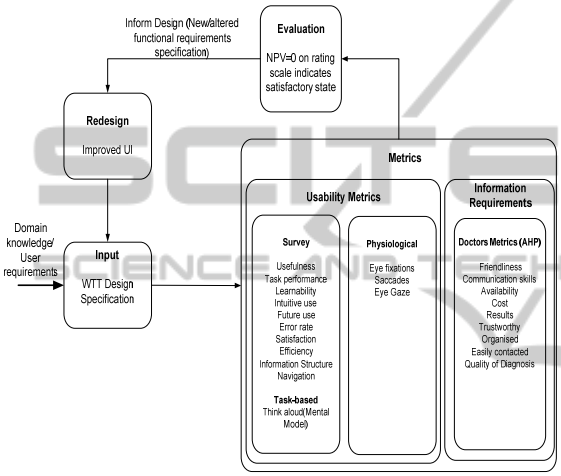

Figure 1: Methodological framework.

3.1 Identifying and Assessing the Importance of Doctors Selection Criteria

Selecting a doctor is a process that requires the evaluation of a number of criteria. These could be different for each patient but in large numbers there is a convergence in the most important issues that all patients are interested in. The identification of these criteria was performed using a combination of Analytical Hierarchy Process (AHP) (Saaty, 1990) with domain knowledge from the literature (Powis, 2003). For this activity 20 participants were involved. Each participant should have been under at least one medical treatment. During the AHP process, each participant was asked to specify their selection criteria. In addition, researchers conducted a literature review to identify further aspects relating to criteria selection in medicine. The set of criteria that emerged was then given to the participants to prioritise. This was performed using a pairwise comparison of each criterion with each other using a numerical scale from 9 to 1/9. Results from each evaluation yielded the specification of an evaluation matrix. The Eigen value and consistency values of the matrix are subsequently calculated, so as to eliminate conflicting opinions. An accumulated average of each raw in the matrix gave rise to the weighting of each criterion. The weighs were then normalised in a rage of 0-100 and the result was assigned to each rating question in the doctors' evaluation page on the website. Based on the ratings of each user the score of each doctor was calculated using the weight assigned to each rating question.

3.2 Eye Tracking and Thinking Aloud in Usability Testing

Eye tracking is a widely used method for detecting usability problems (Nielsen et al, 2010). Eye tracking provides information regarding users' visual attention. This can be expressed in eye fixations, when participant fix their eyes on an object, or saccades, which denote the movements between fixations. However, research by Hyrskykari et.al. (2008) state that eye tracking data cannot be always clear for interpretation. Participants might spend some time looking at an object either because it is interesting or confusing. To that end, eye tracking is usually combined with data from other usability testing methods. One of these techniques is thinking aloud which enables the researcher to extract additional task-related information that is participants' head-mental models. Hyrskykari et.al.(2008) describe two ways to perform the thinking aloud method. One of them is to ask users to explain what they are doing/thinking during the task i.e. concurrent think aloud (CTA). The other is to ask participants to verbalize their thoughts after each task or after all tasks are completed, i.e. retrospective think aloud (RTA). When it task performance measurement is of essence, it is better to use RTA since some people might get overwhelmed if they perform the task and think aloud at the same time.

3.2.1 Experimental Usability Evaluation

The objective of the test was to evaluate the website whattheythink.info: a new website created for providing and collecting information and reviews about all doctors in Cyprus. To determine the behaviour of participants, it is necessary to observe users during their interaction with the experimental conditions (Haynes et al. 2003). Therefore, researchers observed users' interaction with the website through an eye tracking technology to identify what they

found interesting, confusing, or irrelevant to their task. In addition, researchers measured task completion time, task success, errors rate and user satisfaction. During the experiment researcher were taking notes regarding participants' behaviour along with their externalised thoughts from think aloud protocol. After the experiment users had to complete the post-test questionnaire.

4 METHODOLOGY

The evaluation method used initiated with the design of the experiment, the specification of the questionnaires and the identification of the evaluation criteria from the literature. The main constructs of the questionnaire were: trust, satisfaction, usability and functionality. The experiment followed a usability testing approach. Initially a group of participants that satisfied the initial criteria were selected. A pilot study was conducted to test the reliability (Cronbach's alpha) and validity of the questionnaire and helped to identify problems with the experiment. The experiment was conducted to assess the usability of WTT using eye tracking, thinking aloud and survey techniques. Also the experiment addressed issues relating to task completion times, errors and success rate. Throughout the experiments, participants' interactions were monitored and video-recorded for further analysis. Think-aloud information was mapped on a timeline with eye tracking patterns and tasks in the experiment (Figure 2). This helped to associate physiological behaviours with mental activities and phenotype behaviours. Use of a temporal analysis method to map each phenotype behaviour to mental model state helped to identify the search strategy of each participant and accordingly find dominant patterns among strategies.

With the completion of the experiment the data collected were pre-processed and analysed to identify patterns of behaviour that highlighted areas of concern. The problems that have been identified where used for the redesign of the website.

Fifteen participants took part in the study, 7 male aged 20 to 45. All of them were students or professors from European University Cyprus. Three tasks were given to each participant and performance was recorded for each one. Each test session took on average 20 minutes. During the experiment with the eye tracker, participants were asked to look on the screen until the end of the experiment to be able to record their eye fixations with the eye tracker. Users had to perform three tasks. In the case that the user didn't know how to complete a task, the task was skipped and was considered failed. During their interaction with the system, users were asked to think aloud. This help the researcher expose the mental models of the participants. Specifically, subjects were expected to describe what they were thinking while executing a task. This helped to verbalise the thinking processes of participants in the case of errors or confusion.

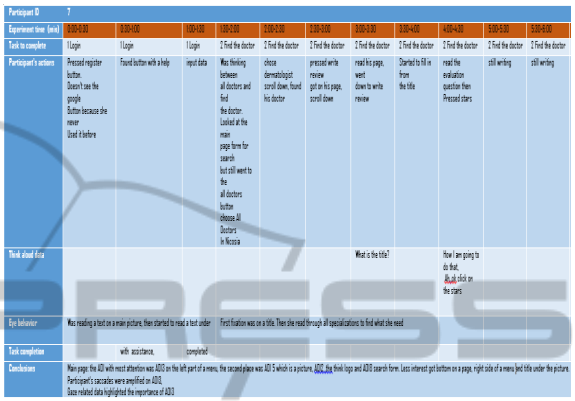

Figure 2: The temporal analysis of the think aloud experimental data and eye behaviour for the three tasks, for one participant. The timeline is shown at the top of the table and at each time interval the relevant observations regarding think aloud and eye tracking, are reported in the cells underneath.

The experiment was conducted in a usability testing lab. For the physiological observations, researchers used the Tobii eye tracker along with video recording equipment. Pre and post questionnaires were used to get the user evaluations on the constructs shown in figure 1. Before the experiments, users were asked to fill in a demographic questionnaire. Next, all participants had to perform 3 tasks during the experimental process. The scenario tasks were designed in a way to help the researchers identify areas of concern. The three tasks in the scenarios included: registering to the website using either, manual, Gmail or LinkedIn, Facebook methods. The second task requested from the participants to identify a dentist that they had recently visited and write a review of their experience. The final task was to identify the most highly rated Cardiologist for a particular city. Throughout the experiment the subjects were encourages to think aloud. The researchers were monitoring the interaction of the subjects with the system and were recording performance data such as errors and task completion times. All sessions were video-recorded for the transcription of the think aloud data. After the experiment all participants were asked to fill in a user satisfaction, usability and trust questionnaires. Participants had to give

answers on a 7-point Liker scale questions. The questionnaire consisted of 2 parts. The first part included 19 questions covering issues relating to ease of use, satisfaction and trust. The second part of the questionnaire consists of 15 open-ended questions that aim to collect user opinions regarding the functionality and usefulness of the website.

4.1 Eye Tracking with Think Aloud

According to Nielsen and Pernice (2010), the mind-eye-behaviour explains that what people are looking at and what they are thinking about tends to be the same. Eye tracking can usually give a quite clear view of what people are paying attention to. Usually users look at the same thing they are thinking about. This leads to the conclusion that fixations (when the eye is resting on something) equal attention. Interpreting eye behaviours has been a challenge to researchers in the field. A comprehensive summary of the interpretations of eye behaviours is reported by Ehmke and Wilson (2007)

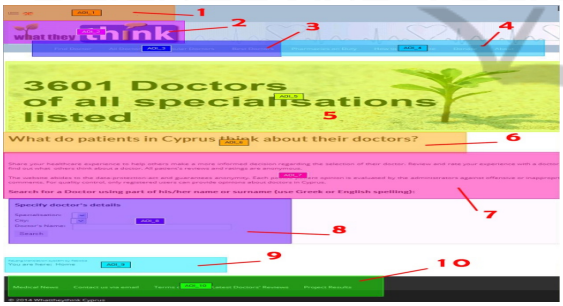

Figure 3: Specified AOIs for the main page of the site.

The interpretation of the results from the experiments was based on eye-movement metrics and related usability problems as reported in the literature (Ehmke and Wilson 2007). However, pure content of eyes fixation is not enough as researchers also would like to know users' feelings when they look at a certain part of a website; whether the user is happy, satisfied, confused or angry. According to Nielsen and Pernice (2010), users can overlook something because it is not interesting for them or completely unclear. Therefore, the combination of eye tracking with cognitive approaches such as think aloud is necessary to make an accurate interpretation of users' behaviour in evaluation studies. Therefore, during the experimental stage of WTT evaluation data regarding the eye behaviour of participants' were mapped with their mental models in a temporal table. The table shows in a timeline the activity of each user (task) and their mental process (think aloud). Each task and mental activity is associated with a time-stamp. In the same way data from the eye tracker is recorded on a timeline. Therefore, the mapping between participants' eye behaviour with the task and mental activities was possible by using as a reference point the time-stamp of each critical event in both timelines. Eye tracking data was interpreted based on eye-movement relating to fixations, saccades, scan-path and gaze (Ehmke and Wilson, 2007). Results from this were administered in a temporal table shown in Figure 2.

5 DATA ANALYSIS

During the experiment, Tobii Studio was used to record eye behaviour of participants. This aimed in monitoring user attention. An aggregated view of all participants' attention yielded heatmaps for each page of the site. High concentration areas on the heatmap are colored in red. The areas of the, that received less interest is colored in yellow, and areas with the least-user attention in green. The rest of the website that has no colour didn't get any user attention. Figure 3 depicts the participants' heatmap for the main page of WTT. To evaluate the level of attention per areas of interest (AOI), the website's pages were divided into a number of AOIs. This enabled the assessment of the distribution of eye fixations between participants and AOI. Figure 3 also depicts the AOIs and the distribution of user fixations among them, for the main page of the site (Figure 4).

Figure 4: The distribution of eye gaze per AOIs and the heat map of participants' attention (below) for the main page of the site.

Results from the eye behaviour analysis helped to identify problems in the information architecture

of WTT. Specifically, data from this activity helped to evaluate the current placement of information on the site, the labelling systems such as, information representation and the choice of appropriate terminology considering target audience. Moreover, eye tracking data helped to assess WTT's navigation systems, plus, analyse if users can easily find relevant information. For each participant, eye tracking data were integrated with think aloud data using the method described in Figure 2. Hence, recorded interpreted eye behaviour patterns with think aloud key findings on a temporal chart. An aggregated analysis of all charts highlighted issues related to the current design of WTT. In addition, qualitative data collected from the open questions of the interview process, was analysed using sentiment analysis. The classification was based on positive and negative sentiment towards WTT. Overall the results showed a positive opinion towards the use of the site. However, some issues were reported regarding the doctor search functionality. In terms of task completion rate, 90% of participants completed task1, 80% task 2and 70% task 3. Most errors were reported during task 2, while searching for the right doctor and in task 3 while rating the selected doctor. Addition data collected, from participants in the post and pre-experiment questionnaires. These addresses questions relating to three dimensions of the website, mainly trust, usability and usefulness. Results from this analysis was assesses using the Net Positive Value method (NPV) to indicate the positive effect of each question and hence to easily pin point areas of concern. The scale was 7 point scale with a 1 (very poor) to 7 (very good) to increase the discrimination in the evaluators' judgement. They were asked to report the reasons for their decisions and any interaction problems they had observed under the relevant heuristic. These rating scores were converted into net positive values (NPV) to reflect the range of the users' assessments. Figure 5 shows the NPV results for the post experiment questionnaires.

Table 1: Assessment of the three dimensions of WTT.

Category	Score
Ease of use	4.2/7.0 (Adequate)
Satisfaction	4.9/7.0 (Good)
Trust	4.8/7.0 (Good)

5.1 System Redesign

The study helped to identify problems with the existing design of the website. This drove the requirements refinement process (Figure 6) that took into consideration usability guidelines. Subsequently, the refined requirements were used to redesign the website.Table 1 illustrates the average score for each of the three constructs that were evaluated. This indicates that the main concern in WTT is ease of use and hence the WTT redesign focused on that. In the same vein, an improved usability of a website should have a positive effect on trust (Seckler et al, 2015).

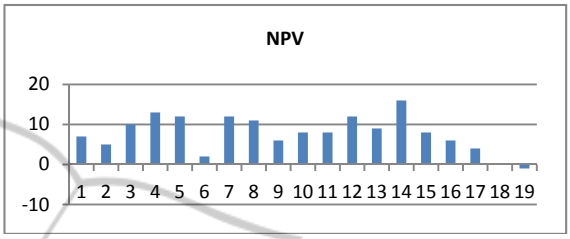

Figure 5: WTT NPV scores.

Figure 6: The redesigned site.

During the experimental evaluation of WTT, researchers found several problems with the searching and labelling of information. The solution for the labelling problem was to remove unnecessary navigation buttons and improve the wording of labels based on a revised analysis of target users' characteristics. Moreover, the doctors search issue was due to the naming of doctors in the database. Doctors were not found when spelling of the users was not identical to that used in the database. This problem was resolved using partial specification of doctors' name in the search criteria. In addition the search space is reduced by setting partial search criteria, i.e. the city where the doctor practises to increase the probability of finding a match. Moreover, to eliminate the problem with the information search, additional contextual cues relevant to the search task were introduced through an improved information architecture. The new interface of the website after taking into consideration the results of the evaluation is depicted below.

6 CONCLUSIONS

The main goal of this research was to devise a method for evaluating doctors review websites such as the WTT. A literature review was conducted to identify influencing factors that affect online patients' decision making process, trust and satisfaction. The underlying research method used in this study was design science and the artifact was the WTT website. The focus was on the evaluation of the artifact and the definition of redesign specification to overcome identified problems. The evaluation of the artifact combined techniques such as, eye tracking, surveys and think aloud to enhance the accuracy of the results. Physiological and cognitive data from the experiments were integrated on a temporal scale that helped to pinpoint problems in a holistic way.

Results from the evaluation of WTT indicated an average usability score. Moreover, analysis of the integrated data obtained from eye tracking, video recording, and 'thinking aloud' methods, also highlighted problems with the website's information architecture. Both led to the need to redesign the site. Despite the fact that all problems that were identified have been eliminated, the redesigned website also needs to be re-evaluated. Therefore, part of our future work, is to perform a comparative study between the two designs using a larger sample set. Furthermore, we also intent to enhance our methodology with additional behavioural cues such as stress level of users (Carneiro D, et al, 2012). This will help identify stressors in the information architecture that lead to reduced artifact acceptance.

REFERENCES

Berben, S. A., Engelen, L. J., Schoonhoven, L. and Van De Belt, T. H. (2010). Definition of health 2.0 and medicine 2.0: a systematic review. *Journal of medical Internet research*, 12(2).

Brezillon, P., Carlsson, S., Respicio, A. and Wren. P. (2014), *DSS 2.0 – Supporting Decision Making With New Technologies*, IOS press.

Carneiro. D., Castillo. J. C., Novais. P., Fernández-Caballero. A. and Neves. J. (2012). Multimodal Behavioural, Analysis for Non-invasive Stress Detection, *Expert Systems With Applications*, 39(18), 13376-13389.

Chatterjee, S. and Hevner, A. (2010). *Design Research in Information Systems*. New York: Springer US.

Davis, M. M., Gebremariam, A., Hanauer, D. A., Singer, D. C. and Zheng, K. (2014). Public Awareness, Perception, and Use of Online Physician Rating Sites, *JAMA*, 311(7).

Davis, F. D., Davis, G. B., Morris, M. G. and Venkatesh, V. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3).

Eisenfuhr, F. (2011). *Decision making*. New York, NY: Springer.

Ehmke, C. and Wilson, S. (2007). *Identifying Web Usability Problems from Eye-Tracking Data*. Retrieved from. http://bcs.org/upload/pdf/ewic_hc07_lppaper12.pdf.

Fox, S. (2011, May). The social life of health information. *The Pew Internet & American Life Project*.

Gabbard, J. L., Hix, D, and Swan, E. J. (1999). User Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*, 19(6), 51-59.Galizzi M., Miraldo M., Stavropoulou C., Desai, M., Jayatunga, W., Joshi, M. and Parikh, S. (2012). Who is more likely to use doctor-rating websites, and why? *BMJ Open*, 2(6).

Hannon, N. S., Lagu, T., Lindenauer, P. K. and Rothberg, M. B. (2010). Patients' evaluations of health care providers in the era of social networking. *J Gen Intern Med.*, 25(9), 942-946.

Haynes, S. N., Heiby, E. M. and Hersen, M. (2003), *Comprehensive Handbook of Psychological Assessment*. New Jersey: John Wiley and Sons.

Howe, J. (2006), *The Rise of Crowdsourcing*. Retrieved http://archive.wired.com/wired/archive/14.06/crowds.html.

Hyrskykari, A., Ovaska, S., Majaranta, P., Räihä, K□J. and Lehtinen, M. (2008). Gaze path stimulation in retrospective think aloud. *Journal of Eye Movement Research*, (Vol. 2).

Maxwell, K. (2002). The Maturation of HCI: Moving Beyond Usability toward Holistic Interaction. in Carroll, J.M. (ed.), *Human-Computer Interaction in the New Millennium (*pp. 191-209). New York: Addison-Wesley.

Morand, D. A. and Ocker, R. J. (2002). Exploring the Mediating Effect of Group Development on Satisfaction in a Virtual and Mixed-Mode Environment. *e-Service Journal*, (Vol. 1, pp. 25-42).

Nielsen, J. and Pernice, K. (2010). *Eyetracking Web Usability*. Berkeley: New Riders.

Nielsen, J. (1996). Usability Metrics: tracking interface improvements. *IEEE Software,* (Vol. 13).

Riley, S. (2012). *Social Networks For Doctors Aid Medical Crowdsourcing*. Retrieved August 8, 2012, from http://news.investors.com/technology/081412-622143-medical-doctors-communicate-via-own-social-networks.htm?p=full.

Powis, D. (2003). Selecting medical students. *Medical Education*, 37:1064-1065.

Saaty, T. L. (1990). How to make a decision: The Analytic Hierarchy Process. *European Journal of Operational Research,* 48, 9−26.

Seckler. M, Heinz. S., Forde. S., Tuch. A. and Opwis. K. (2015). Trust and distrust on the web: User experiences and website characteristics. *Computers in human behaviour*, 45, 39 -50.