Enriching Second-Screen Experiences with Automatic Content Recognition

Jorge Abreu[1], Pedro Almeida[2], Telmo Silva[3]

[1] University of Aveiro - Digimedia, Portugal, jfa@ua.pt

[2] University of Aveiro - Digimedia, Portugal, almeida@ua.pt

[3] University of Aveiro - Digimedia, Portugal, tsilva@ua.pt

Keywords: Second-screen, Interactive Television, User Experience, Usability, Evaluation methodologies.

Abstract. Technological devices surrounding the television are changing, leading to changes in viewers' habits and to the development of second-screen applications created to provide better TV viewing experiences. Used while watching television, the 2NDVISION application presented in this paper is able to identify – through audio-fingerprint – content being displayed on the TV screen and present enhanced information on the second screen. Under a participatory design approach, the development of the application took in consideration the users' opinion regarding its main functionalities and interface solutions and included evaluation test sessions conducted in laboratory settings. This paper reports on the first evaluation sessions and the preliminary results gathered through a cognitive walk-through methodology, with opinions regarding instrumental, non-instrumental and emotional impact of the application collected through SUS, AttrakDiff and SAM scales.
The results show that users experienced no major navigation problems and part of the iconography was validated. Considering the main goals of the application users were satisfied and interested in having such an application for providing additional information about the TV shows they watch.

Introduction

As technological devices surrounding the television changed, so have viewers' habits, with the general TV experience no longer narrowed to a single and fixed location (typically a shared space at home) but extending to new locations and integrating different platforms. Mobile devices (smartphones and tablets) are being used not only as additional platforms for watching video contributing for the "space shifting" phenomenon [1] but also as companion devices, allowing for a greater involvement with what is being watched on (a regular) TV – e. g. by performing general web searches or to search for additional information related with the program being watched [2].

This is a consequence and simultaneously a stimulus for the focus of the Interactive Television (iTV) industry on the development of second screen applications, designed to deliver supplementary information related with the TV content [3]. In fact, and mainly due to the ubiquitous character of mobile technologies, the opportunity for well-designed mobile applications and services to influence positively on individuals' day lives increases [4]. Nevertheless, and for this to happen, it is necessary to look beyond the product's usability and start to consider the impact of emotions and aesthetics in the overall impression of the system.

In this paper, the authors report on the 2NDVISION companion App to be used while watching TV being able to automatically identify content displayed on the television screen and presenting related synchronized (using audio fingerprint) information. Conducted under a participatory design approach, the development of the application took into consideration the users' opinion towards functionalities, aesthetic and appraisal of the system.

After establishing the theoretical background (in section two of the paper), section three introduces the application main features and system architecture and addresses the methodological approaches that were taken during the development process, namely the first set of tests carried to validate a preliminary prototype in three dimensions – instrumental qualities, non-instrumental qualities and emotional reactions. Finally, section four presents the research results; and section five presents the preliminary conclusions.

State of the art

Second-screen applications

The growing success of second-screen devices is changing the way users relate and interact with the television. Using second-screen devices while watching TV is an increasingly common activity: according to recent numbers of Nielsen Company [5], 62% of North Americans and 44% of European consumers used second screen devices while watching TV. Consumers are increasingly adopting a lean forward approach to the television experience, using connected devices as extensions of the program they are watching [5].

Secondary Screen applications are, in the context of 2NDVISION project and according to the description of Red Bee Media [6], those that provide a companion experience, increasing and improving synchronously the viewer experience with content related to that he is watching in the TV. By synchronizing the App with television content, it is possible to provide enhanced-information (e.g. biographical data on a given actor, the name of the song that is playing, more details on the narrative of a series or film) or trigger events correlated with what is happening on TV, particularly in the context of interactive advertising.
Existing mobile apps perform synchronization with the TV content essentially based on audio fingerprint systems. This technology, which has collected the public recognition through applications such as Shazam (http://www.shazam.com), allows audio monitoring regardless of its format and without the need for meta-data or watermarks [7]. It is therefore possible to associate certain "signatures" or patterns to the audio meta-information providing or triggering events correlated with what the user is watching on TV. Given the increasing processing power of mobile devices, the evolution of algorithms and cloud-computing techniques, it is possible to use techniques of image content recognition that, in essence, allow, on the one hand, to compensate the inherent limitations in the audio fingerprint and on the other, to open up new possibilities and use cases in the field of the secondary screen apps.

In this context, there has been the development of various applications, in particular for the US market, that exploits the synchronization capabilities with content to be displayed on TV. Again Shazam is an example, but also applications such as Viggle or Beamly. However, despite the above context, the specific offer for the Portuguese market is not available and 2NDVISION may contribute to enrich the Portuguese TV ecosystems with an end to end system allowing the app to synchronize (using audio fingerprint) and enrich both linear and non-linear (e.g. Catch-up TV or VoD) TV content.

UX in second-screen scenarios

The specificities of mobile devices created numerous and significant challenges in the field of user experience. Mobile context, multimodality, connectivity, small screen size, different display resolutions and power emerged as factors to be taken into consideration when designing interfaces for mobile devices[8]. Interacting with these devices implies a different look at user experience. User's hands are no longer over the mouse but directly interacting with the interface through multi-touch gestures such as swipe, pinch, press and hold [9], [10], creating the need for bigger buttons (to solve the "fat-finger" problem), wider distance between icons and new navigation paradigms.

Mobile usability models often focus on the effectiveness, efficiency and satisfaction, disregarding the cognitive load and the emotional impact of interacting with applications in no-longer defined time and place. When developing second-screen applications designed to enhance the TV viewing experience, it is even more delicate to take into account dimensions such as cognitive capacity and attention selectiveness.

The 2NDVISION App

The growing adoption of second-screen devices while watching TV emphasized the importance of developing solutions able to balance the user's attention between two or more sources of information (e.g. the TV screen and the mobile phone or tablet screen), namely by creating applications able to deliver additional content in a ease and user-friendly way. When developing solutions towards the user's engagement in the TV viewing experience, designers must take into consideration the pragmatic and non-pragmatic dimensions of the system. While developing second-screen apps, as it is imperative to achieve a balance between the information in the first and in the second screen and to design solutions for minimizing attention dispersion [11], usability is a key issue. Nevertheless, it is also important to consider the user's overall appraisal of the application [12], making it imperative to include them and consider their opinion during the design and development of the product.

Based in audio fingerprint and image recognition technology, the "2nd-screen" Android application is able to automatically identify content displayed on the TV and present related synchronized information (hereafter defined as markers, i.e. pieces of information composed by images, text and links for related sources of information); to automatically aggregate all markers in a sequential index; to trigger notifications whenever new markers are displayed. Users could also save (bookmark), rate, filter and share markers; configure notifications to be triggered when new content is detected; and select programs to follow.

System architecture

The 2NDVISION is based on an Event-Driven Architecture (EDA), which meets high performance in real-time events, immediate action to the consumers, "fire & forget" integration and a quality of service with the shortest waiting time possible.

This architecture is composed by a series of modules (depicted in Figure 1), namely the Automatic Content Recognition (ACR) which gears the second screen device with the ability to perceive the content which is being visualized; via the indexed markers by the ACR module, the solution has a Recommendation Engine that is responsible for the creation of additional contents. An Audience Engager module is responsible for obtaining user profiles through their interaction with the solution; and through the Advertising and Campaign Planning Engine it becomes possible to manage campaigns and advertisements related to the indexed content. In order to guarantee the persistent and in cache content management, the Content Repository module uses relational and non-relational systems for data

management. The module responsible for the configuration of the solution, match requests processing between the app and content, markers configuration, image processing, users profile configuration and statistical reports generation is the Back Office. It is also possible, via the Gamification Engine to attribute a gamming environment to the 2NDVISION app and generate quizzes, polls and leaderboards, among other forms of user engagement. Finally, the Client Gateway is the communication channel between the APIs and allows the interaction between the mobile app and the backend of the solution.

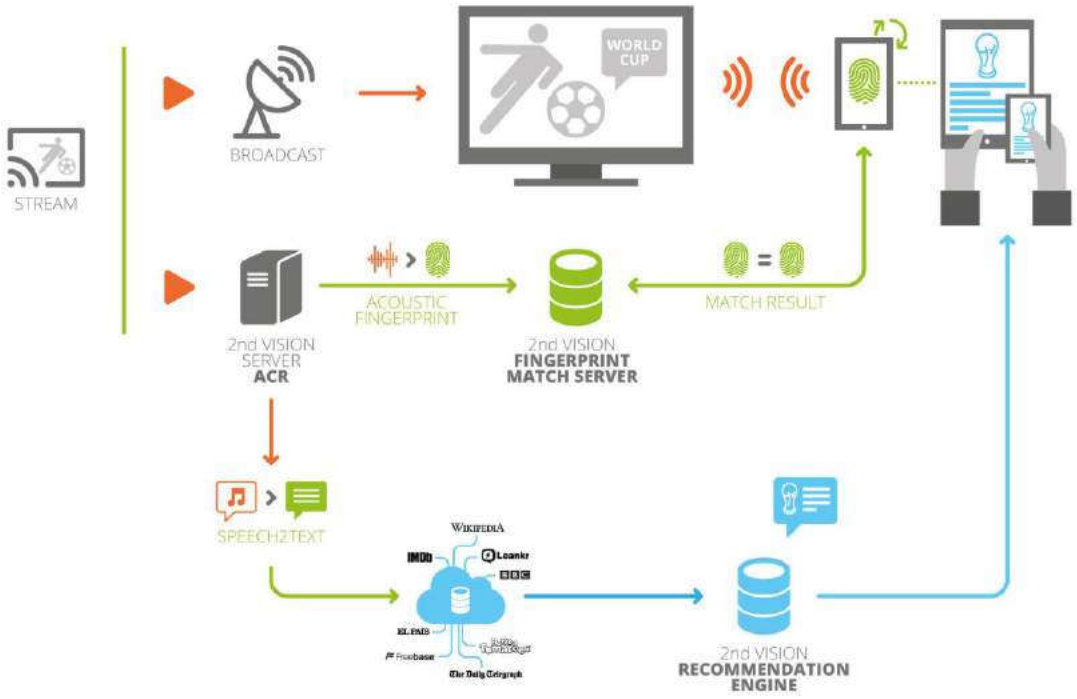

Figure 1 - 2NDVISION system architecture

Features

The 2NDVISION system architecture enables offering its users a wide range of features. This allows the structuring of the application in several areas. The following figure (Figure 2) identifies 6 of the 8 areas/features available in the 2NDVISION App.

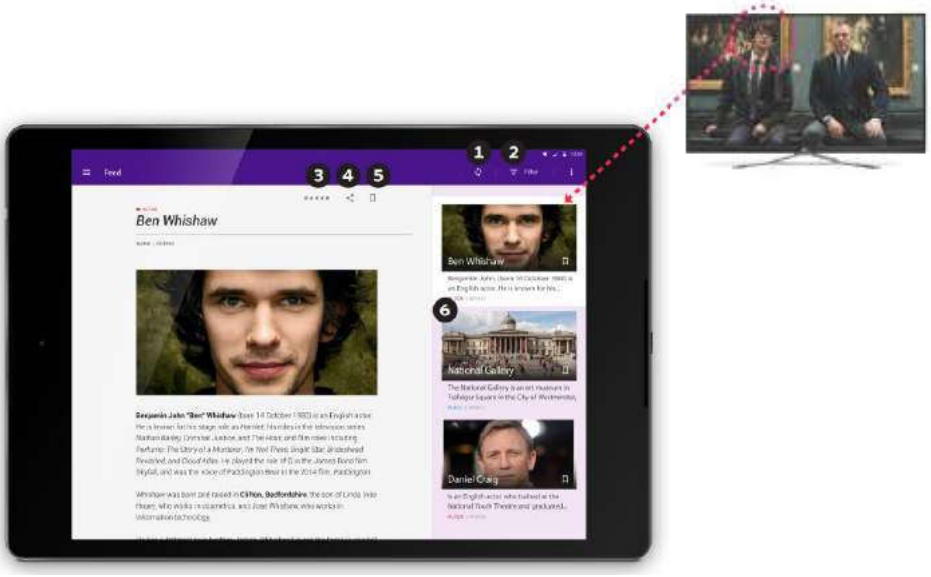

Figure 2 - The 2NDVISION interface - main screen

1. **Synchronize:** Triggering the automatic identification of the TV show being aired provides related and synchronized markers;
2. **Filter**: Content filtering categories. Example: Places (location, architecture, scenic spots, weather);
3. **Rate:** User's classification of the additional information displayed;
4. **Share:** Sharing any marker via social networks and e-mail;
5. **Bookmark:** Archiving of detected content for later viewing;
6. **Feed:** Area that displays the additional information related with the TV program (markers).

Evaluation

Preliminary evaluation

As it was established that all the application design and development would be conducted under a participatory design approach, the research team carried test-sessions aiming to evaluate users' perception about the applications instrumental and non-instrumental qualities, as well as their emotional reactions to the system. The first evaluation session, reported in this paper, was carried in March 2015. The evaluation was carried by introducing evaluators with a first prototype of the 2NDVISION application (Figure 3). The prototype allowed users to explore the most relevant features that were designed for the application.

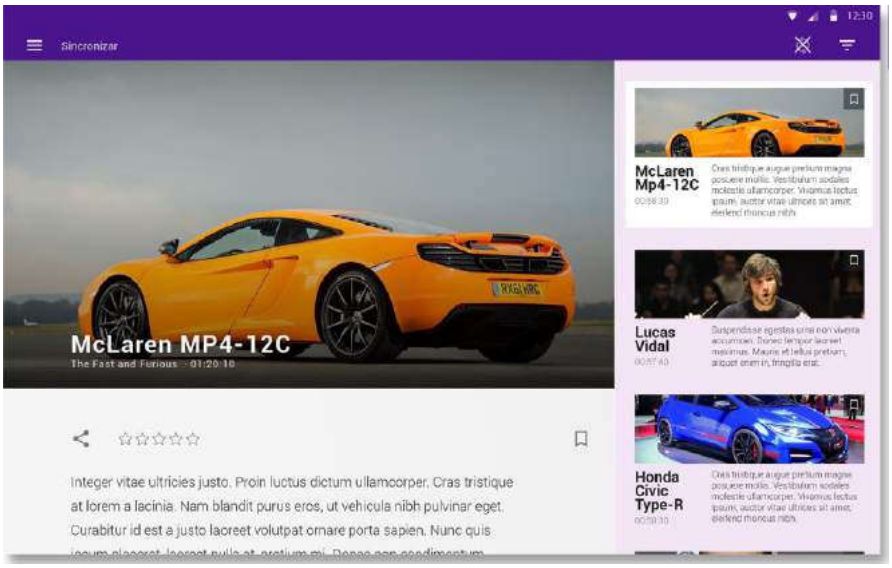

Figure 3 – The prototype used in the first evaluation session

Tests sessions were conducted individually in laboratory settings replicating a regular living room. The main goals of the tests were: i) to evaluate the graphical interface; ii) to validate the consistency of the adopted solutions (icons, layout and content organization); iii) to identify critical interface issues; iv) to evaluate the Non-instrumental qualities, and; v) the emotional impact on potential users.

At the beginning of the test, a print with the application' main features was handed-out to the participant. The print had two illustrations of the application's main screen, indicating its main features: automatic synchronization of content with additional information displaying. Each participant was left alone to freely explore the application during 2 minutes, after which

he/she was informed about the structure of the test session and introduced to the application functionalities. Participants were then invited to perform a set of tasks mainly focused on navigation (e.g. navigating between markers, exploring the index); iconography (e.g. correspondence between the icon and the related action); and in the use of specific features (e.g. saving, rating and sharing markers, scheduling programs in order to receive notifications). Considering that the application was still in the prototype phase and therefore some features were not yet fully operational, the tests were focused on the validation of the interface, namely in what concerns to navigation, iconography and specific features.

During the test, each participant had to perform 6 tasks centred on iconography, navigation and implementation of specific operations. In the end, a short interview was conducted and each participant was asked to answer three questionnaires.

All tasks were performed under a cognitive walkthrough approach, and all data was recorder. After performing the tasks, and taking the CUE-model [12] as a framework and following adaptations already mentioned in literature review (cf. ([13]; [14];[15]; [16]), three instruments were used in order to assess the users' perspectives about the applications' instrumental and non-instrumental qualities, and the user's emotional reactions to the whole experience.

For assessing the applications' instrumental qualities (e.g. controllability, effectiveness, learnability), the research team used the SUS - System Usability Scale [17] and the Pragmatic dimension of AttrakDiff [18]. The SUS-scale [17] is a 10-item questionnaire with a 5-point Likert scale developed to assess several usability aspects such as ease of use and usefulness. As for AttrakDiff [18], it is an instrument created to measure the attractiveness of an application by presenting the user/evaluator three groups of opposite adjectives. AttrakDiff records both the perceived pragmatic quality, the hedonic quality and the attractiveness of an interactive product or application.

For accessing non-instrumental qualities (e.g. aesthetics, identification) the research team used the AttrakDiff Hedonic dimension.

Finally, for the evaluation of emotional reactions (e.g. valence, arousal, control of the application), the research team used the SAM - Self-assessment Manikin [19]. The SAM-Manikin is a non-verbal pictorial assessment method that directly assesses the pleasure, arousal, and dominance associated with the user's affective reactions to a certain stimuli.

Qualitative data was also collected through a semi-structured interview conducted at the end of the test session, aiming to get the participant's opinion regarding the application (its main features), the interface and the adopted solutions.

Participants

Eighteen evaluators (10 men and 8 women) between 18 and 25 years old joined the sessions. With regard to qualifications, 9 were college students, 3 graduates, 4 had a master degree and 2 had a doctoral degree. With regard to the profession/occupation, the participants were distributed among students (5), programmers (4), researchers (3) and other professions.

Of the 18 participants, 16 used to watch TV programs, 13 had already used applications to search for TV programs related information and 16 were regular users of social networks. Each session lasted an average of 10 minutes.

Instrumental qualities

Instrumental qualities were assessed through the SUS scale and the Pragmatic dimension (PQ) of the AttrakDiff instrument. In the SUS scale scores go from 0 to 100, with scores under 50 indicating that the assessed product is not acceptable in terms of usability; from 60 to 70 the acceptance is marginal; from 71 to 100, products are considered as acceptable. In AttrakDiff, scores go from -3 to 3, and the closer the score is to “3” the better it is in terms of usability. Data analysis shows that the prototype scored 77,1 (in a maximum of 100) in the SUS scale, and 1,1 (in a maximum of 3) in the Pragmatic dimension of AttrakDiff. This means that the application had a good level of usability [20] but there was room for improvements in what concerns its pragmatic qualities. This conclusion is consistent with the data collected through interviews: when asked to share their opinion regarding the application and its main functionalities, participants pointed that it was an “interesting” application, namely by its ability to automatically present content related with TV programs. Regarding the application’s interface, most participants described it as “easy to use”, “intuitive”, “interesting”, “accessible” and “well organized”. Nevertheless this global opinion, participants also mentioned some points to be improved: according to their opinion, some icons (e.g. the index categories filter button) were not clear about its function, and others (e.g. the “share” and “bookmark” buttons) were considered small. According to some participants, some feedback messages were missing. 3 out of the 18 participants mentioned that although considering the application interesting, they would not use it because they were not regular TV consumers.

Non-instrumental qualities

The non-instrumental qualities of the application (related to aesthetics and connection with the application) were evaluated through the hedonic-identity (HQ-I) and hedonic-stimulation (HQ-S) quality dimensions of AttrakDiff. In AttrakDiff, the HQ-I classification reflects the extent to which the user is identified with the product, while the HQ-S classification reflects the degree of user satisfaction in terms of novelty, interest and stimulus. In these dimensions, the application scored 0,9 (HQ-I) and 0,8 (HQ-S) (in a maximum of 3), indicating that the implementation satisfied the conventional standards in terms of aesthetic attractiveness and contributed to user identification with the product, but there is room for improvement. These findings are consistent with the qualitative feedback received in the interviews conducted at the end of the test sessions. As noted above, participants indicated that some icons were too small and/or were not sufficiently clear regarding their function.

Emotional Impact

Emotional reactions of the participants were collected by applying the SAM-Manikin scale, which evaluates the emotional reaction of users in three dimensions: satisfaction, motivation, and sense of control. On this scale, the score ranges between 1 and 9 and the results can be considered negative if they are between 1 and 4, the neutral value is "5", and positive if the classification obtained is between 6 and 9. In this scale, the application got a rating of 7,3 on the satisfaction dimension; 6,1 on what concerns motivation; while sense of control was rated as 7,1. Although the application needs to be improved in the motivation dimension, the remaining values indicate that there are no major problems in what concerns the emotional impact of the use of the application.

Conclusions and recommendations

During the tests, in an overall perspective, no major problems were revealed in terms of navigation (e.g. access to previous markers, access to other features/menu areas); and part of the iconography has been validated.
According to the evaluators feedback, it was possible to identify the following suggestions for improvement: i) to make it clear for the user that he may navigate between markers on the main page by including the image of the beginning of the next marker image side by side to the main marker; adjust in size and shape some icons; include additional configuration and personalization settings.

This first evaluation was very important in the development of the interface and interaction options. The conclusions from the first evaluation session were integrated in the following versions of the prototype and a final evaluation setting was prepared.

Considering the main goals of the application users were satisfied and interested in having such an application in providing additional information about the TV shows they watch. These are some important results that motivate the team to conclude the development of the 2ndVision App.

Current and future work includes the conclusion of the App, the development of a new lab evaluation and preparing a pilot (field trial). It is expected that users have the full user experience in the field trial and the generated data will allow the development team to define on the potential market opportunity for such an application. Concerning the evaluation results, the researchers expect to compare the results gathered in the different development stages of the application to validate the evaluation methodology and compare the first motivation and expectation results with the final field trial results.

Acknowledgements

This paper is a result of the 2ndVision project, funded by QREN (grant agreement no. 38783). Authors are grateful to Mónica Aresta, Lígia Duro, Rita Oliveira and André Ferreira for their work on the project and also to the remaining project partners: Altran and Outsoft.

References

[1] Jancovich, Mark. 2011. Time, Scheduling and Cinema-going. Media International Australia, Incorporating Culture & Policy, No. 139, May 2011: 88-95. Availability:<http://search.informit.com.au/documentSummary;dn=055670613543388;res=IELLCC> ISSN: 1329-878X.

[2] Abreu, J., Almeida, P., Teles, B., and Reis, M. 2013. Viewer behaviors and practices in the (new) television environment. In Proceedings of the 11th european conference on Interactive TV and video (EuroITV '13). ACM, New York, NY, USA, 5-12. DOI=http://dx.doi.org/10.1145/2465958.2465970

[3] Geerts, D., Leenheer, D., Heijstraten, S. & Negenman, J. 2014. In Front of and Behind the Second Screen: Viewer and Producer Perspectives on a Companion App. In Proceedings of the TVX'14 Conference, (95-102). June 25 - 27, 2014, Newcastle Upon Tyne, United Kingdom

[4] Sun, X. & May, M. 2014. Design of the User Experience for Personalized Mobile Services. International Journal of Human Computer Interaction (IJHCI), Volume (5): Issue (2)

[5] The Nielsen Company. 2015. Screen wars: the battle for eye space in a TV-everywhere world, March 2015. Retrieved June 13, 2015, from http://www.nielsen.com/us/en/insights/reports/2015/screen-wars-the-battle-for-eye-space-in-a-tv-everywhere-world.html

[6] Red Bee Media. 2012. Second Screen Series - Paper 1: Setting The Scene. Retrieved from http://www.redbeemedia.com/sites/all/files/downloads/second_screen_series_paper_1_whitepaper_red_bee_media.pdf

[7] Cano, P., Batle, E., Kalker, T., & Haitsma, J. 2002. A review of algorithms for audio fingerprinting. Paper presented at the 2002 IEEE Workshop on Multimedia Signal Processing

[8] Zhang, D, & Adipat, B. 2005. "Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications". International Journal of Human-Computer Interaction, 18:3, 293-308, DOI: 10.1207/s15327590ijhc1803_3

[9] Treder, M., Pachucki, A., Zielonko,A. & Łukasiewicz, K. 2014. "Mobile book of trends 2014". UX Pin & Movade internal report, http://studio.uxpin.com/ebooks/mobile-design-book-of-trends/

[10] Bank, Chris. & Zuberi, W. 2014. "Mobile UI Design Patterns". UX Pin & Movade internal report, http://studio.uxpin.com/ebooks/mobile-design-patterns/

[11] Lee, K., Flinn, J., Giuli, T., Noble, B. & Peplin, C. 2013. "AMC: Verifying User Interface Properties for Vehicular Applications". MobiSys'13, June 25-28, Taipei, Taiwan

[12] Mahlke, S. & Thuring, M. 2007. "Studying Antecedents of Emotional Experiences in Interactive Contexts". CHI 2007 Proceedings - Emotion & Empathy. San Jose, CA

[13] Gross, A. & Bongartz, S. 2012. "Why do I like it?: investigating the product-specificity of user experience". In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (NordiCHI '12). ACM, New York, NY, USA, 322-330. http://doi.acm.org/10.1145/2399016.2399067

[14] Law, E., Schaik, P., & Roto, V. 2014. "Attitudes towards user experience (UX) measurement". International Journal of Human-Computer Studies, Volume 72, Issue 6, Pages 526-541, ISSN 1071-5819, http://dx.doi.org/10.1016/j.ijhcs.2013.09.006.

[15] Bach, C., Gauducheau, N. & Salembier, P. 2011. "Combining interviews and scales in the multidimensional evaluation of user experience: a case study in 3D games". In Proceedings of the 29th Annual European Conference on Cognitive Ergonomics (ECCE '11). ACM, New York, NY, USA, 157-160. DOI=10.1145/2074712.2074743

[16] Aranyi, G., van Schaik, P. 2014. "Modeling user-experience with news Web sites". Journal of the Association for Information Science and Technology. 1-23. http://onlinelibrary.wiley.com/doi/10.1002/asi.23348/epdf

[17] Brooke, J. 1996. "SUS-A quick and dirty usability scale". Usability evaluation in industry, PW Jordan, B. Weerdmeester, A. Thomas and IL McLelland, 189-194

[18] Attrakdiff. 2011. http://www.attrakdiff.de/

[19] Bradley, M. M., & Lang, P. J. 1994. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), 49-59

[20] Bangor, A., T. Kortum, P., & T. Miller, J. 2008. An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, 24(6), 574- 594