Multimed Tools Appl (2017) 76:16125–16143
DOI 10.1007/s11042-016-3898-9

# Users' viewpoint of usability and user experience testing procedure - gaining methodological insights in a case of an interactive HbbTV application

**Jože Guna**[1] · **Emilija Stojmenova-Duh**[1] · **Matevž Pogačnik**[1]

Received: 16 March 2015 / Revised: 9 June 2016 / Accepted: 24 August 2016 /
Published online: 15 September 2016


**Abstract** We present a novel meta-methodological approach for the user experience and usability methodology procedure evaluation, shown in an example of the user experience and usability study of an interactive HbbTV application. The idea behind this research is not only to evaluate and improve the TV-WEB service but also to gain insights how the participants perceived the whole Ux evaluation procedure itself. The research questions focused mainly on the time complexity (temporal demand) and the frustration level of the TV-WEB evaluation procedure. Additionally the appropriateness of the selected content used, interface and interaction design, and the service impressions and satisfaction/payment related questions was sought. A special questionnaire based on the NASA TLX standard test is presented. The concept has been successfully implemented in several live field trials in three countries (BiH, Serbia and Montenegro). In the user experience and usability study of the service itself more than 150 participants were involved, of those 35 took part in the meta-methodology study. The feedback was quantitatively evaluated on a 7-point Likert scale, with "1" indicating the best positive feedback, "4" neutral/undecided feedback and "7" the most negative feedback. The quantitative average summary results obtained in the three evaluation studies were 1.30 for both the BiH and Serbia test study cases and the score of 1.22 for the test study case in Montenegro. These results show that a great majority of the participants found the whole evaluation procedure time-wise and frustration-wise undemanding, with appropriate content, presentation style and overall attitude towards them. By using this approach it was possible to improve the user experience and usability methodology used, producing more reliable results and providing better user experience in the final version of the product as well as providing a pleasant experience during the testing of the product.

**Keywords** TV and Internet · User study · User experience · User centered design · Human-computer interaction · Meta-methodology · SEE TV-WEB project

✉ Jože Guna
joze.guna@fe.uni-lj.si; http://www.ltfe.org

[1] Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1000 Ljubljana, Slovenia

## 1 Introduction and motivation

With the rapid evolution of multimedia and Internet technologies the evolution of networked multimedia platforms followed in the form of convergence towards digital information representation and conveyance. Digital television is the manifestation of the trend, evolving from the passive and rigid medium towards a fully interactive environment supporting a plethora of innovative service schemes. The possibilities of digital television extend far beyond the rigid service space of analogue television by enabling the combination of audio-visual information to interactive services traditionally perceived as being foreign to television. Digital television and the belonging interactive multimedia rich services have the potential to revolutionize the way in which the general population perceives and deals with television [10].

Television has been changing in multiple dimensions: from analogue to digital; from scheduled broadcasts to on-demand television on the internet; from a lean-back (passive) to a lean-forward (active) media; from straight watching to the consumption of content connected to additional services; from the sole television viewer to the viewer being part in social networks and communities regarding to the television content, etc. Two of these trends, which are especially apparent, are social and collaborative interactive television in convergence with the internet as the information/communication/collaboration media. Television was a social media from the very beginning, where people gathered to enjoy a football match together, were watching the first steps of mankind on the moon or relaxed in front of an exciting movie. The actual "collaboration" took place outside the media, e.g. in follow-up face-to-face conversations. With the development of more advanced media formats on the web, interactivity emerged. The TV environment as such is therefore a changing environment, which more and more adapts ideas of media services as known from the web, or using the web parallel as information/communication/collaboration media [2].

Nevertheless, usability and the overall user experience (Ux) with a product or a service, especially in a case of the multimedia and interaction rich TV applications, may prove to be a deciding factor of success or failure. Therefore, designing a service, targeting a broad and diverse audience, including users who potentially have no previous experience with technology or even show aversion toward it, is challenging, at the least. Beside the requirements for the technology–wise mature solution, effective human-to-human design and communication, human/user centered design, the user interface, interaction method, usability and user experience in general, play a paramount role in the final service's success and acceptance. To minimize these risks the users should be involved at all stages of the development process. This approach is commonly referenced by the term User-Centered Design (UCD) [11]. The final goal of the UCD approach is not only to develop a useful service (usability aspect), but also to provide a service that is easy and enjoyable to use with a low cognitive demand and low learning curve, thus ensuring a good user experience. Usability is usually denoted as the ability of the user to use the device or service in order to successfully carry out a task at hand, whereas the user experience takes a broader view, looking at the users' entire interaction with the device or service, as well as the thoughts, feelings and perceptions that result from that interaction [3]. The importance of these topics in conjunction with the semantic ambient media, now and in the future, is also discussed and emphasized in [12], ultimately leading to a more efficient, ecological, safe, and human friendly future.

Choosing the right methodology and measuring the user experience [1] is not a trivial task, as it dependents on the product or service itself, whilst being heavily influenced by the target user group. To illustrate these aspects, we present a meta-methodological approach towards the

evaluation of user experience and usability methodology within the context of an interactive HbbTV television application. The study is based on a real life case, an international EU SEE TV-WEB project [16]. The aim of the TV-WEB project is to use the free digital terrestrial television (DTT) broadcasting frequency spectrum capacities for transmitting selected Internet content, and to provide the best possible experience for the elderly people, economically weak people, and people living in rural areas, who do not usually use personal computers and who have no possibility of an Internet connection at all. The concept of the project differs from the services provided by technologies, such as connected and Smart TV, where the Internet experience is ensured by connecting the television set to the Internet. Instead, the SEE TV-WEB project foresees delivery of Internet content to the homes solely using the DTT spectrum. The project involves 16 partners from 7 countries: Austria, Hungary, Slovenia, Croatia, Bosnia and Herzegovina, Montenegro and Serbia. The pilot implementation plan and consequently user experience testing plan was ambitious as well, as there were multiple live pilots performed within all partner countries. There were 21 live pilot system presentations in all 7 partner countries, where more than 350 participants were directly included in the user experience and usability testing and more than 100.000 users were informed about the project goals indirectly [14]. Overall impression of the TV-WEB service was excellent, where more than 80 % of users expressed genuine interest in the service, and would use it on a daily basis. Of those, more than 190 of users were 50 years or older, where additional care was taken to make the whole experience as non-demanding and pleasant as possible.

Considering all these facts, within the project's activities a special attention was given to the most appropriate user experience and usability methodology technique selection, and the overall evaluation procedure itself. In addition to this, a very special care was given to the evaluation of the user experience and usability methodology used. A special questionnaire was developed to better understand these aspects and to possibly improve the user experience and usability methodology used, producing more reliable results and providing better user experience in the final version of the product as well as providing a pleasant experience during the testing of the product.

The main contributions of the present paper are therefore:

- a meta-methodological approach for the user experience and usability methodology procedure evaluation based on a modified NASA TLX standard methodology questionnaire,
- results and insights gained from the meta-methodological evaluation.

The rest of the paper is organized as follows: related work is presented in section 2; a brief description of the SEE TV-WEB project and the user experience methodology and evaluation procedure used is given in section 3; the proposed meta-methodological approach for the user experience and usability methodology procedure evaluation is presented in section 4; results are presented and discussed in sections 5 and 6 respectively; while key conclusions and future work references are drawn in the last section.

## 2 Related work

Designing the user experience for the smartTV environment is different from designing for the PC or mobile devices and presents a different set of challenges. In "User Interface Design Principles for Interactive Television Applications" [4], the authors propose set of user interface design principles specifically targeted for the iTV applications.

The SEE TV-WEB project has already been presented to the public on several professional and scientific conferences and events. In the paper "WEB on TV: designing the user experience" [7] an introduction, project motivation, technical requirements and architecture, and initial usability aspects and proposed user evaluation method for the TV-WEB project are presented.

Specifically addressing the user experience challenges within the TV-WEB project the authors of the "WEB on TV: Cross-National User Study for Designing Better User Experience" [15] research paper report the user studies about the media literacy, media use and the interest about such services in two different European countries, i.e. Austria and Montenegro. Additionally, the user-centered design process for the service that was conducted as a part of the study in Slovenia is presented. Interviews, focus groups and card sorting techniques were used as methodology. Results of the studies were used as initial input for deciding on the most appropriate content and design, and interaction approaches when implementing the TV-WEB service.

Work in the present paper is a continuation of that described in [8], where a novel meta-methodological approach for the user experience and usability methodology procedure evaluation was presented, shown in an example of the user experience and usability study of an interactive HbbTV application. A special questionnaire, partly based on the NASA Task Load Index (NASA TLX) standard test was used.

NASA-TLX [9] is a multi-dimensional questionnaire designed to obtain workload estimates in several dimensions, namely mental, physical and temporal. It enables to obtain a measure of the effort required, performance and overall frustration level from one or more operators while they are conducting a specific task or immediately afterwards. Over the past 20 years, the tool has proven to be reasonably easy to use and reliably sensitive to experimentally important manipulations. Its use has spread far beyond its original application (aviation), focus (crew complement), and language (English) and has been applied to several other domains, also proving to be a valuable source of insight for user experience and usability related applications. An example of a very specific domain use is the Driving Activity Load Index (DALI), an adaptation of the NASA TLX for assessing the mental workload in driving activities [13].

The results in the user study presented in this paper add a small, but important piece to the puzzle. Focusing not only on the evaluation of the service, but also on the end users, their needs and subjective perception of the evaluation procedure itself, provides valuable additional insight how best to design effective and user friendly evaluations. The concept has been successfully implemented in several field trials of the SEE TV-WEB project and very positive feedback was gained, which is further discussed in this work.

## 3 The SEE TV-WEB project and evaluation method procedure

The SEE TV-WEB project's main focus is on providing selected web based multimedia content directly over the DVB-T/T2 networks [6] with simulated local interactivity. The TV-WEB application is built according to the HbbTV standard requirements [5]. The technical solution of the project consists of the server side and the end-user side. The server side provides all the necessary equipment for ensuring content editing, storage, adaptation, multiplexing and distribution over the DVB T/T2 channel, while this content is received and presented to the user on the end-user's side. All information, consisting of the application and multimedia data itself is transmitted directly over the DTT network, rendering the Internet network access unnecessary.

From the users' perspective, the TV-WEB service represents "just another TV channel", in form of a locally interactive TV application. Content is organized into a matrix of news items. The left column represents content types (e.g. local news, weather …), whereas rows present specific news items relevant for each content type. A typical TV-WEB service main screen is shown in Fig. 1.

Special care was given to the most appropriate user experience and usability methodology technique selection, target user group selection and persona modelling, and scenario/tasks/procedure preparation. Methodology approaches, such as think-aloud protocol in combination with a guided interview and observation of verbal as well as non-verbal responses were found to be most appropriate. In addition to this and to ensure the repeatability and comparability of different tests and to help better guide the process, a special questionnaire (Table 1) was developed. The questionnaire served as guidance as it was used by the Ux experts to gain the desired information. The participants did not have to fill out the questionnaire by themselves, as this was done by the Ux experts. Special target user groups, such as the elderly or technologically inexperienced participants, proved to be especially challenging.

The TV-WEB service received a lot of positive feedback, from all target user groups, and interestingly, even the younger participants found value in the service, mainly in its simplicity of use. Some visual impressions from tests performed are show in Fig. 2.

## 4 The TV-WEB methodology procedure evaluation

In addition to the Ux procedure itself, a very special care was given to the evaluation of the user experience and usability methodology used. A special questionnaire (Table 2) was developed to better understand these aspects.

The idea behind this research is not only to evaluate and improve the TV-WEB service but also to gain insights how the participants perceived the whole Ux evaluation procedure itself. This step was also performed in a form of a guided interview after the TV-WEB Ux evaluation procedure was complete. It was designed to be as effortless as possible for the participants and was therefore estimated to be done in 5 to 10 min. The participation in this evaluation was completely optional. The approach is partly based on the NASA Task Load Index (TLX)

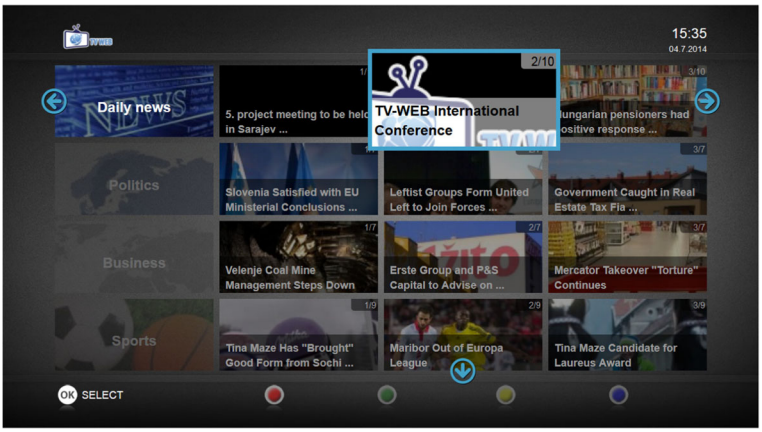


**Fig. 1** The TV-WEB service main screen

**Table 1** Detailed TV-WEB Ux evaluation questionnaire structure

| **TV-WEB Ux questionnaire** | | | |
|---|---|---|---|
| Location | | Date | |
| **1. Personal information** | | | |
| Gender [F, M, or multiple answers in case of a group evaluation] | | | |
| Age | | | |
| Use of technology [Yes, No] | | | |
| Technology experience (1=no experience; 5=regular user)<br>TV: 1 2 3 4 5<br>Computer: 1 2 3 4 5<br>Internet&applications: 1 2 3 4 5 | | | |
| Internet access [Yes, No] | | | |
| Additional notes | | | |
| **2. Tasks** | | | |
| Task 1 | What is the number of news stories in the category »X«? | | |
| | Task completed: | Successfully | Unsuccessfully |
| | Problems: | Yes | No |
| | Notes: | | |
| Task 2 | Search for the news titled »X« and find out who/what is »Y«. | | |
| Task 3 | Set your optimal font size. | | |
| **3. User interface interaction, navigation and design** | | | |
| Do you have any problems using the remote control? | | | |
| Is the navigation clear and easy to use? | | | |
| Are the “breadcrumbs” clear enough? | | | |
| Do you find the functionality of the coloured buttons useful? | | | |
| Would you like to have a special key on the remote control to exit the application? | | | |
| Would you like to have a special key on the remote control that will navigate you directly to the home screen? | | | |
| Is the font size large enough? | | | |
| Is the font easy to read? | | | |
| Do you like the design? Is the contrast ratio good enough? | | | |
| **4. Content** | | | |
| Is the selected content appropriate? Would you like to add/change something? | | | |
| **5. Overall impression and satisfaction** | | | |
| Do you find the service simple to use? | | | |
| Would you like to use the service daily? | | | |
| If yes, how much would you be willing to pay for the purchase of the service (both hardware and software)? | | | |
| Overall impression and satisfaction with the service. Comments, wishes. | | | |

**Fig. 2** Impressions from the TV-WEB Ux evaluation (Podgorica, Montenegro, June 2014)

**Table 2** The TV-WEB methodology evaluation questionnaire

| **TV-WEB methodology evaluation questionnaire** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Location | | Date | | | | | | |
| **Personal information** | | | | | | | | |
| Gender [F, M, or multiple answers in case of a group evaluation] | | | | | | | | |
| Age | | | | | | | | |
| Use of technology [Yes, No] | | | | | | | | |
| Technology experience (1=no experience; 5=regular user)<br>TV: 1 2 3 4 5<br>Computer: 1 2 3 4 5<br>Internet&applications: 1 2 3 4 5 | | | | | | | | |
| Internet access [Yes, No] | | | | | | | | |
| Additional notes | | | | | | | | |
| **N** | **Question** | **Positive** | | **Undecided** | | | | **Negative** |
| **1** | **Physiological-psychological aspects** | | | | | | | |
| 1.1 | Physical difficulty/effort of solving the TV-WEB Ux procedure (simple vs. extremely complex) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1.2 | Mental difficulty/effort of solving the TV-WEB Ux procedure (simple vs. extremely complex) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1.3 | Time complexity of solving the TV-WEB Ux procedure (brief vs. extremely time consuming) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1.4 | Presentation style – procedure/task/question clarity (clear/understood vs. extremely confusing) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1.5 | Presentation style - attitude towards participants (pleasant vs. extremely unpleasant) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 1.6 | Presentation style – ambient setup (pleasant vs. extremely unpleasant) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **2** | **Procedure content/task selection** | | | | | | | |
| 2.1 | Personal information section appropriateness (appropriate vs. not appropriate) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.2 | Task selection overall appropriateness (appropriate vs. not appropriate) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.3 | Task 1: What is the number of news stories in the category »X«? | | | | | | | |
| 2.3.1 | Task 1 – time complexity (brief vs. extremely time consuming) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.3.2 | Task 1 – intuitiveness, pleasantness (intuitive/pleasant vs. not intuitive at all/frustrating) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.4 | Task 2: Search for the news titled »X« and find out who/what is »Y«. | | | | | | | |
| 2.4.1 | Task 2 – time complexity (brief vs. extremely time consuming) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.4.2 | Task 2 – intuitiveness, pleasantness (intuitive/pleasant vs. not intuitive at all/frustrating) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.5 | Task 3: Set your optimal font size. | | | | | | | |
| 2.5.1 | Task 3 – time complexity (brief vs. extremely time consuming) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.5.2 | Task 3 – intuitiveness, pleasantness (intuitive/pleasant vs. not intuitive at all/frustrating) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| | | | | | | | | |
| 2.6 | Content related questions (appropriate vs. not appropriate) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.7 | User interface interaction, navigation and design related questions (appropriate vs. not appropriate) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| 2.8 | Overall service impressions and satisfaction/payment related questions (appropriate vs. not appropriate) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **3** | **Overall grade/satisfaction** | | | | | | | |
| 3.1 | Time spent while solving the TV-WEB Ux procedure (subjective feeling of time spent in minutes) | < 10 | 10 - 20 | 20 - 30 | 30 - 40 | 40 - 50 | 50 - 60 | > 60 |
| 3.2 | The TV-WEB Ux procedure overall grade - frustration? (not frustrating/pleasant vs. extremely frustrating) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| **Additional notes:** | | | | | | | | |

standard test [8, 9] and partly custom designed with the TV-WEB specifics in mind. This approach makes it useful for the intended purpose while still allowing comparison with other standardized approaches and studies.

The NASA TLX is a very simple, yet effective tool for assessing the work load. Originally designed to better understand the physical and cognitive level of difficulty of astronauts' task, the test can be effectively applied to other domains as well. The standard test consists of 6 questions using the 7-point Likert scale, as shown in Fig. 3. Mental, physical, temporal, performance, effort and frustration aspects for a specific task can be measured, making the tool valuable for user experience measurement as well.

The research questions focused mainly on the time complexity (temporal demand) and the frustration level of the TV-WEB evaluation procedure. The aim of the research questions was to gain answers on the physical, mental, time complexity and frustration level of the participants with the evaluation procedure as a whole, as well as to gain insights of the time complexity and frustration level of the specific user tasks (Table 1) in the procedure. Additionally the appropriateness of the selected content used, interface and interaction design, and the service impressions and satisfaction/payment related questions was sought.

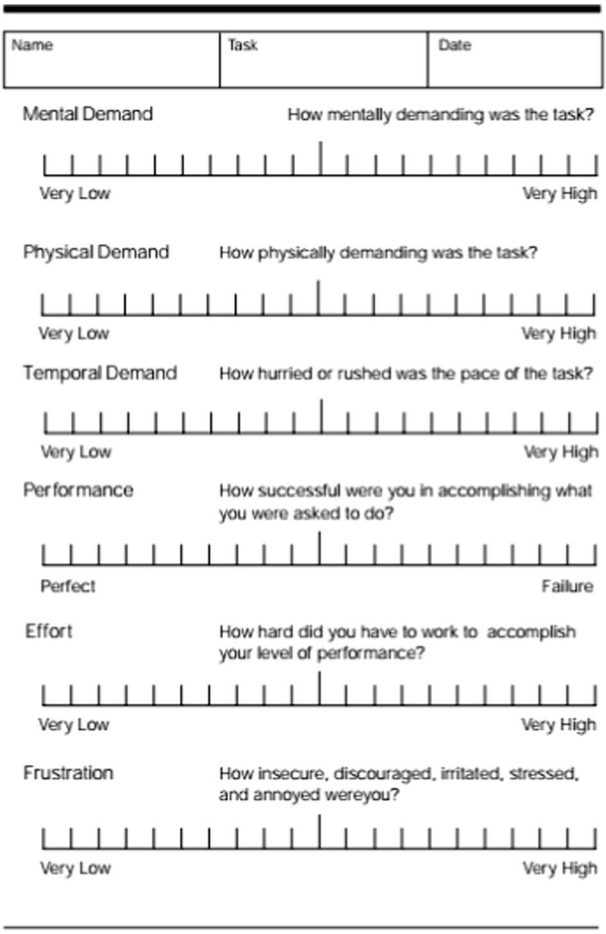


**Fig. 3** The NASA TLX standard workload test

The main research questions were:

- What is the perceived time complexity (temporal demand) of the evaluation procedure as a whole?
- What is the perceived frustration level of the evaluation procedure as a whole?
- What is the perceived time complexity (temporal demand) of the task 1 of the evaluation procedure?
- What is the perceived frustration level of the task 1 of the evaluation procedure?
- What is the perceived time complexity (temporal demand) of the task 2 of the evaluation procedure?
- What is the perceived frustration level of the task 2 of the evaluation procedure?
- What is the perceived time complexity (temporal demand) of the task 3 of the evaluation procedure?
- What is the perceived frustration level of the task 3 of the evaluation procedure?
- Are the selected content related questions used in the evaluation appropriate?
- Are the user interface interaction, navigation and design related questions appropriate?
- Are the overall service impressions and satisfaction/payment related questions appropriate?

A seven point Likert scale (the same as in the NASA TLX) was used to evaluate these questions, with "1" indicating the best positive feedback, "4" neutral/undecided feedback and "7" the most negative feedback. The metrics used is the average score attainted per each question per test trial location and the average score of all results per test trial location. Lower scores represent a more positive result.

The personal and previous technology experience information is gained from the Ux test questionnaire (Table 1). The rest of the evaluation questionnaire (Table 2) is divided into three parts, corresponding with the content of the Ux test questionnaire and procedure from Table 1.

The questions regarding the mental, physical, temporal, performance, effort and frustration aspects are based on the NASA TLX test (1.1, 1.2, 1.3, 2.3.1, 2.3.2, 2.4.1, 2.4.2, 2.5.1, 2.5.2, 3.1 and 3.2) while the questions regarding the presentation style, attitude towards the participants and appropriateness of the content are TV-WEB project related (1.4, 1.5, 1.6, 2.1, 2.2, 2.6, 2.7 and 2.8) and based on the previously selected TV-WEB service evaluation methodology (Table 1).

Three specific tasks were included in the TV-WEB evaluation procedure. In the first task (What is the number of news stories in the category »X«?) the participants had to navigate the TV-WEB interface and count a number of news items/stories under a specific category. In the second task (Search for the news titled »X« and find out who/what is »Y«.) the participants had to open a specific news item/story and find a specific information in the text. In the third task (Set your optimal font size.) the participants had to use the colour button functionality and set an optimal font-size. In the proposed methodology evaluation procedure the aim was to determine the time complexity of these tasks and the perceived frustration levels.

The proposed methodology evaluation procedure was used in several live field test trial of the TV-WEB project, namely in Serbia, Montenegro and Bosnia. Tests were performed in:

- Sarajevo, Bosnia and Herzegovina in April 2014,
- Novi Sad, Belgrade and Čačak, Serbia in May 2014, and
- Podgorica, Montenegro in June 2014.

## 5 Results

In Bosnia two pilot presentations and usability tests were held. The first one at the faculty of Electrical Engineering in Sarajevo and the other one in a private house in a small rural town near Sarajevo. In all, 46 participants were included, with genders equally represented.

**Table 3** Detailed user methodology feedback summary results from the test case study in Sarajevo, Bosnia and Herzegovina

| N | Question (number of responses) | Positive | | Undecided | | | | Negative | AVG | STD |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **Physiological-psychological aspects** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 1.1 | Physical difficulty/effort of solving the TV-WEB Ux procedure | 13 | | 1 | | | | | **1.14** | **0.52** |
| 1.2 | Mental difficulty/effort of solving the TV-WEB Ux procedure | 13 | 1 | | | | | | **1.07** | **0.26** |
| 1.3 | Time complexity of solving the TV-WEB Ux procedure | 8 | 6 | | | | | | **1.43** | **0.49** |
| 1.4 | Presentation style – procedure/task/question clarity | 14 | | | | | | | **1** | **0** |
| 1.5 | Presentation style - attitude towards participants | 14 | | | | | | | **1** | **0** |
| 1.6 | Presentation style – ambient setup | 13 | 1 | | | | | | **1.07** | **0.26** |
| **2** | **Procedure content/task selection** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 2.1 | Personal information section appropriateness | 12 | 1 | 1 | | | | | **1.21** | **0.56** |
| 2.2 | Task selection overall appropriateness | 14 | | | | | | | **1** | **0** |
| 2.3 | Task 1: What is the number of news stories in the category »X«? | | | | | | | | | |
| 2.3.1 | Task 1 – time complexity | 7 | 5 | 2 | | | | | **1.64** | **0.72** |
| 2.3.2 | Task 1 – intuitiveness, pleasantness | 8 | 5 | 1 | | | | | **1.50** | **0.63** |
| 2.4 | Task 2: Search for the news titled »X« and find out who/what is »Y«. | | | | | | | | | |
| 2.4.1 | Task 2 – time complexity | 5 | 7 | 2 | | | | | **1.79** | **0.67** |
| 2.4.2 | Task 2 – intuitiveness, pleasantness | 7 | 5 | 2 | | | | | **1.64** | **0.72** |
| 2.5 | Task 3: Set your optimal font size. | | | | | | | | | |
| 2.5.1 | Task 3 – time complexity | 10 | 4 | | | | | | **1.29** | **0.45** |
| 2.5.2 | Task 3 – intuitiveness, pleasantness | 10 | 4 | | | | | | **1.29** | **0.45** |
| | | | | | | | | | | |
| 2.6 | Content related questions | 14 | | | | | | | **1** | **0** |
| 2.7 | User interface interaction, navigation and design related questions | 12 | 1 | 1 | | | | | **1.21** | **0.56** |
| 2.8 | Overall service impressions and satisfaction/payment related questions | 12 | 1 | | 1 | | | | **1.29** | **0.80** |
| **3** | **Overall grade/satisfaction** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| | **time (minutes)** | **<10** | **10-20** | **20-30** | **30-40** | **40-50** | **50-60** | **>60** | | |
| 3.1 | Time spent while solving the TV-WEB Ux procedure | 5 | 6 | 3 | | | | | **1.86** | **0.74** |
| 3.2 | The TV-WEB Ux procedure overall grade - frustration? | 10 | 2 | 1 | | | | | **1.31** | **0.61** |
| **Summary: 14 participants; average summary answer score = 1.30** | | | | | | | | | | |

Approximately ¾ of them were familiar with computers and had some Internet experience. Approximately 50 % of users were 20–50 years old, 3 % less than 20 years and the rest over 50 years. Of those, 14 volunteered to participate in the meta-methodology study, as well.

Four pilot presentations were held in Serbia. The first pilot was carried out at the Technology Fair 2014 in Belgrade, the second one in a daily center for elderly citizens in Belgrade, a further presentation at the faculty of Technical Sciences in Čačak and finally the

**Table 4** Detailed user methodology feedback summary results from the test case study in Novi Sad, Serbia

| N | Question (number of responses) | Positive | | | Undecided | | | Negative | AVG | STD |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **Physiological-psychological aspects** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 1.1 | Physical difficulty/effort of solving the TV-WEB Ux procedure | 6 | | | | 1 | | | **1.57** | **1.40** |
| 1.2 | Mental difficulty/effort of solving the TV-WEB Ux procedure | 5 | 1 | | 1 | | | | **1.57** | **1.05** |
| 1.3 | Time complexity of solving the TV-WEB Ux procedure | 5 | 1 | | 1 | | | | **1.57** | **1.05** |
| 1.4 | Presentation style – procedure/task/question clarity | 6 | 1 | | | | | | **1.14** | **0.35** |
| 1.5 | Presentation style - attitude towards participants | 7 | | | | | | | **1** | **0** |
| 1.6 | Presentation style – ambient setup | 7 | | | | | | | **1** | **0** |
| **2** | **Procedure content/task selection** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 2.1 | Personal information section appropriateness | 3 | | | | | | | **1** | **0** |
| 2.2 | Task selection overall appropriateness | 3 | | | | | | | **1** | **0** |
| 2.3 | Task 1: What is the number of news stories in the category »X«? | | | | | | | | | |
| 2.3.1 | Task 1 – time complexity | 6 | | | 1 | | | | **1.43** | **1.05** |
| 2.3.2 | Task 1 – intuitiveness, pleasantness | 7 | | | | | | | **1** | **0** |
| 2.4 | Task 2: Search for the news titled »X« and find out who/what is »Y«. | | | | | | | | | |
| 2.4.1 | Task 2 – time complexity | 5 | | | 1 | | | | **1.50** | **1.12** |
| 2.4.2 | Task 2 – intuitiveness, pleasantness | 6 | | 1 | | | | | **1.29** | **0.70** |
| 2.5 | Task 3: Set your optimal font size. | | | | | | | | | |
| 2.5.1 | Task 3 – time complexity | 5 | 1 | | 1 | | | | **1.57** | |
| 2.5.2 | Task 3 – intuitiveness, pleasantness | 6 | | 1 | | | | | **1.29** | |
| | | | | | | | | | | |
| 2.6 | Content related questions | 6 | | 1 | | | | | **1.29** | **0.70** |
| 2.7 | User interface interaction, navigation and design related questions | 6 | | 1 | | | | | **1.29** | **0.70** |
| 2.8 | Overall service impressions and satisfaction/payment related questions | 7 | | | | | | | **1** | **0** |
| **3** | **Overall grade/satisfaction** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| | **time (minutes)** | **<10** | **10-20** | **20-30** | **30-40** | **40-50** | **50-60** | **>60** | | |
| 3.1 | Time spent while solving the TV-WEB Ux procedure | | 4 | | | | | | **2** | **0** |
| 3.2 | The TV-WEB Ux procedure overall grade - frustration? | 6 | 1 | | | | | | **1.14** | **0.35** |
| **Summary: 7 participants; average summary answer score = 1.30** | | | | | | | | | | |

fourth one at the faculty of Technical Sciences in Novi Sad. The service was tested by 79 users, of those 53 % female. Of those, 86 % of them were familiar with computers and had some Internet experience. Approximately 43 % of users were 20–50 years old, 34 % between 50 and 65, and the rest over 65 years. Of those, 7 participants in Novi Sad volunteered to participate in the meta-methodology study, as well.

**Table 5** Detailed user methodology feedback summary results from the test case study in Podgorica, Montenegro

| N | Question (number of responses) | Positive | | | Undecided | | | Negative | AVG | STD |
|---|---|---|---|---|---|---|---|---|---|---|
| **1** | **Physiological-psychological aspects** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 1.1 | Physical difficulty/effort of solving the TV-WEB Ux procedure | 13 | | | | 1 | | | **1.29** | **1.03** |
| 1.2 | Mental difficulty/effort of solving the TV-WEB Ux procedure | 13 | | | | 1 | | | **1.29** | **1.03** |
| 1.3 | Time complexity of solving the TV-WEB Ux procedure | 13 | | | 1 | | | | **1.21** | **0.77** |
| 1.4 | Presentation style – procedure/task/question clarity | 14 | | | | | | | **1** | **0** |
| 1.5 | Presentation style - attitude towards participants | 14 | | | | | | | **1** | **0** |
| 1.6 | Presentation style – ambient setup | 14 | | | | | | | **1** | **0** |
| **2** | **Procedure content/task selection** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| 2.1 | Personal information section appropriateness | 12 | 2 | | | | | | **1.14** | **0.35** |
| 2.2 | Task selection overall appropriateness | 13 | | | 1 | | | | **1.21** | **0.77** |
| 2.3 | Task 1: What is the number of news stories in the category »X«? | | | | | | | | | |
| 2.3.1 | Task 1 – time complexity | 12 | 1 | | | | | | **1.08** | **0.27** |
| 2.3.2 | Task 1 – intuitiveness, pleasantness | 9 | 1 | | | | | | **1.10** | **0.30** |
| 2.4 | Task 2: Search for the news titled »X« and find out who/what is »Y«. | | | | | | | | | |
| 2.4.1 | Task 2 – time complexity | 9 | 2 | 2 | | | | | **1.46** | **0.75** |
| 2.4.2 | Task 2 – intuitiveness, pleasantness | 7 | 2 | | | | | | **1.22** | **0.42** |
| 2.5 | Task 3: Set your optimal font size. | | | | | | | | | |
| 2.5.1 | Task 3 – time complexity | 13 | 1 | | | | | | **1.07** | **0.26** |
| 2.5.2 | Task 3 – intuitiveness, pleasantness | 13 | 1 | | | | | | **1.07** | **0.26** |
| | | | | | | | | | | |
| 2.6 | Content related questions | 12 | 1 | 1 | | | | | **1.21** | **0.56** |
| 2.7 | User interface interaction, navigation and design related questions | 12 | 1 | | | | | | **1.08** | **0.27** |
| 2.8 | Overall service impressions and satisfaction/payment related questions | 9 | 3 | 1 | | | | | **1.38** | **0.62** |
| **3** | **Overall grade/satisfaction** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | | |
| | **time (minutes)** | **<10** | **10-20** | **20-30** | **30-40** | **40-50** | **50-60** | **>60** | | |
| 3.1 | Time spent while solving the TV-WEB Ux procedure | 7 | 5 | 1 | 1 | | | | **1.71** | **0.88** |
| 3.2 | The TV-WEB Ux procedure overall grade - frustration? | 11 | 1 | 1 | | | | 1 | **1.64** | **1.59** |
| **Summary: 14 participants; average summary answer score = 1.22** | | | | | | | | | | |

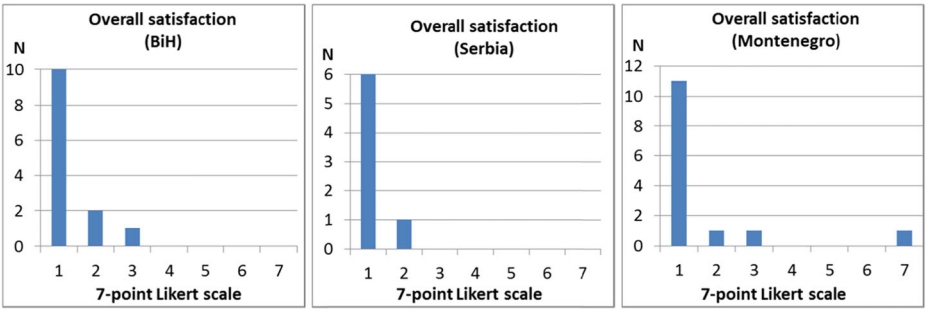


**Fig. 4** The overall satisfaction with the procedure - responses to the question "The TV-WEB Ux procedure overall grade - frustration?", from test study cases in Bosnia and Herzegovina, Serbia and Montenegro (from left to right)

In Montenegro one pilot presentation and usability test was held at the Faculty of Electrical Engineering in Podgorica. The service was tested by the 38 users, 58 % of them male. Approximately ¾ of them were familiar with computers and had some Internet experience. Approximately 26 % of users were 20–50 years old, 34 % less than 20 years and the rest over 50 years. Of those, 14 volunteered to participate in the meta-methodology study, as well.

Detailed results of the meta-methodology evaluation are presented in Tables 3, 4 and 5. The results are in a form of a cumulative number of answers on a 7-point Liker scale under a specific question. The first section is based on the NASA TLX standard methodology test and is used to assess the participants' subjective perception of physical, mental and time related aspects of the TV-WEB evaluation procedure. The second section is directly TV-WEB related and covers interface, content and task related aspects. The last section is used to assess the subjective time spent during the evaluation and the overall impression and frustration levels. The last two aspects are also graphically represented in figures Figs. 4 and 5.

## 6 Discussion

The overall results of the meta-methodological study show a very positive feedback and attitude of the participants toward the Ux and usability study of the TV-WEB service. The meta-methodology evaluation was completely voluntary and supplementary to the main Ux

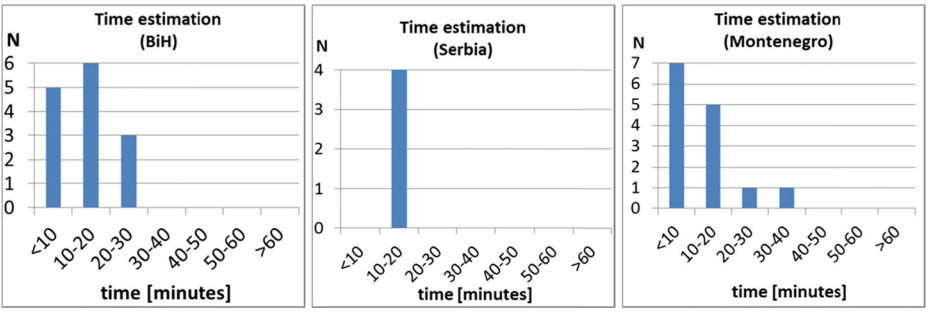


**Fig. 5** The subjective time estimation/perception of the procedure - responses to the question "Time spent while solving the TV-WEB Ux procedure", from test study cases in Bosnia and Herzegovina, Serbia and Montenegro (from left to right)

evaluation survey; therefore fewer participants took part in the survey. All recorded answers are presented in the study. The results of the research questions are further discussed.

## 6.1 The research questions

### *6.1.1 The perceived time complexity (temporal demand) of the evaluation procedure as a whole*

At the test study location in Bosnia and Herzegovina, the perceived time complexity was scored as 1.86, at the test study location in Serbia as 2 and in Montenegro as 1.71. Considering the results, the evaluation session duration was assessed as appropriate and time-wise not demanding. The marginally higher score of 2 in Serbia (indicating a higher time complexity) can be explained due to the few answers received on this research question (only 4 participants).

Objectively, the time allotted for each individual TV-WEB whole Ux evaluation session was up to 1 h per participant. The session dynamics was completely adjusted to the participants in order to provide a stimulating environment and to lessen the potential frustration level. Therefore the sessions took from 10 min to approx. 1 h. Considering the time estimation results, with an approx. average on 10–20 min duration, a great majority of participants deemed the duration of evaluation sessions as appropriate and brief. Only a few time duration impressions were recorded of up to 40 min, which was still completely acceptable considering the initial time limit of 1 h.

### *6.1.2 The perceived frustration level of the evaluation procedure as a whole*

At the test study location in Bosnia and Herzegovina, the perceived frustration level was scored as 1.31, at the test study location in Serbia as 1.14 and in Montenegro as 1.64.

From the results we can conclude that the participants were very satisfied how the evaluation procedure was performed, with only one recorded answer (Montenegro) in all three test study case locations classified as “negative”. The great majority of participants classified the procedure as pleasant and non-frustrating.

### *6.1.3 The perceived time complexity (temporal demand) of the task 1 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived time complexity was scored as 1.64, at the test study location in Serbia as 1.43 and in Montenegro as 1.1. Of all, the most negative was one neutral answer (score 4) recorded in the test study location in Serbia. We can conclude that the task was time-wise found to be not too demanding.

### *6.1.4 The perceived frustration level of the task 1 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived frustration level was scored as 1.5, at the test study location in Serbia as 1 and in Montenegro as 1.08. Of all, the most negative was one answer (score 3) recorded in the test study location in Bosnia and Hercegovina. We can conclude that the task was found to be not too frustrating.

### *6.1.5 The perceived time complexity (temporal demand) of the task 2 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived time complexity was scored as 1.79, at the test study location in Serbia as 1.5 and in Montenegro as 1.46. Of all, the most negative was one neutral answer (score 4) recorded in the test study location in Serbia. We can conclude that the task was time-wise found to be not too demanding.

### *6.1.6 The perceived frustration level of the task 2 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived frustration level was scored as 1.64, at the test study location in Serbia as 1.29 and in Montenegro as 1.22. Of all, the most negative were two answers (score 3) recorded in the test study location in Bosnia and Hercegovina. We can conclude that the task was found to be not too frustrating.

### *6.1.7 The perceived time complexity (temporal demand) of the task 3 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived time complexity was scored as 1.29, at the test study location in Serbia as 1.57 and in Montenegro as 1.07. Of all, the most negative was one neutral answer (score 4) recorded in the test study location in Serbia. We can conclude that the task was time-wise found to be not too demanding.

### *6.1.8 The perceived frustration level of the task 3 of the evaluation procedure*

At the test study location in Bosnia and Herzegovina, the perceived frustration level was scored as 1.29, at the test study location in Serbia as 1.29 and in Montenegro as 1.07. Of all, the most negative was one answer (score 3) recorded in the test study location in Serbia. We can conclude that the task was found to be not too frustrating.

### *6.1.9 Selected content related questions in the evaluation appropriateness*

At the test study location in Bosnia and Herzegovina, the content related questions appropriateness was scored as 1, at the test study location in Serbia as 1.29 and in Montenegro as 1.21. Of all, two most negative answers were recorded (score 3) in test study locations in Serbia and Montenegro. We can conclude that the content related questions and procedure used in the evaluation were appropriate.

### *6.1.10 The user interface interaction, navigation and design related questions appropriateness*

At the test study location in Bosnia and Herzegovina, the interface and design related questions appropriateness was scored as 1.21, at the test study location in Serbia as 1.29 and in Montenegro as 1.08. Of all, two most negative answers were recorded (score 3) in test study locations in Bosnia and Hercegovina and Serbia. We can conclude that the interface and design related questions used in the evaluation were appropriate.

#### *6.1.11 The overall service impressions and satisfaction/payment related questions appropriateness*

At the test study location in Bosnia and Herzegovina, the impressions and satisfaction/payment related questions appropriateness was scored as 1.29, at the test study location in Serbia as 1 and in Montenegro as 1.38. Of all, one most negative neutral answer was recorded (score 4) in test study location in Bosnia and Hercegovina. We can conclude that the impressions and satisfaction/payment interface related questions used in the evaluation were appropriate.

### 6.2 Summary impressions

Considering the results of the first section, a great majority of the participants in all three test study case locations (BiH, Serbia and Montenegro), assessed the Ux procedure as physically, mentally, and time consuming as non-demanding and very positive. The presentation style, ambient setup and the attitude towards the participants was deemed as positive as well. Only a few participants were undecided (answers in range of 3–5 on the Likert scale) in these matters, with no answers on the negative extreme of the scale (answers in range 6–7).

Similar results are found in the second section assessing the interface, content and task related aspects. As the most unproblematic task was found to be the task 3, where the colour button functionality of the user interface was tested in an example of a font size adjustment exercise ("Set your optimal font size."). The participants found this functionality as very welcome and useful in the application. The other two tasks ("What is the number of news stories in the category »X«?", "Search for the news titled »X« and find out who/what is »Y«.") were used to test the navigation aspects of the TV-WEB service and were found to be slightly more difficult. A great majority of users in all three test study case locations found the content used in the application, as well as the navigation concept and visual design as appropriate. Only a few answers recorded were deemed as "undecided" with no answers under the "negative" side of the scale.

The average (average of average individual scores per test study case) score of 1.30 for both test study cases in BiH and Serbia and the score of 1.22 for the test study case in Montenegro was achieved. Considering these quantitative average summary scores it can be concluded qualitatively, that the participants were satisfied with the evaluation approach and the procedure used.

### 6.3 Reflections on the methodology used

The aim of the main user experience study was to assess the interface and interaction design, the readability and content selection and the perceived value of the TV-WEB service. Additionally, the research presented in this paper aimed to answer how time demanding and frustrating for the participants was the whole evaluation procedure itself.

The duration of the sessions typically lasted from 30 min to maximum 1 h, in case the participants were willing to cooperate. In case of the elderly the sessions were shorter and the duration adjusted to the specific participants' needs. The methodology evaluation procedure assessment part was designed to be as effortless as possible for the participants and was therefore estimated to be done in 5 to 10 min.

The main research questions were answered by observing the participants during the tasks, as well as by asking them direct questions and discussing with them on specific topics. To help guide

the study sessions, special questionnaires were developed. The participants did not fill out the questionnaire directly, however. That was done by the Ux team, where two members were present at each session. The Interviewer was conducting the session, while the Observer took notes.

To evaluate the time complexity and frustration level aspects the NASA TLX method questions were applied in part. This was done because of the time constraints in conducting the experiment, where the complete NASA TLX questionnaire could not be used for all aspects sought in the study. Additionally, the study aimed to answer some specific questions regarding the appropriateness of the content, interface and design, and the perceived value of the TV-WEB service. To ensure comparability of the results the same 7-point Likert scale was used to assess these questions.

## 7 Conclusions

In the paper we proposed a new meta-methodological approach to the evaluation of the user experience and usability methodology used. The idea was presented in a real life case – a HbbTV application of the SEE TV-WEB project.

By using this procedure not only the user experience and usability aspects of the tested service or product are revealed, but also valuable insights from the participants' perspective of the whole user experience procedure are gained. The approach was successfully implemented in several field trials and very positive feedback was gained. The feedback was quantitatively evaluated on a 7-point Likert scale, with "1" indicating the best positive feedback, "4" neutral/undecided feedback and "7" the most negative feedback. The quantitative average summary results obtained in the three evaluation studies were 1.30 for the BiH and Serbia test study cases and the score of 1.22 for the test study case in Montenegro. These results show that a great majority of the participants found the whole evaluation procedure as physical, mentally and time-wise undemanding, with appropriate content, presentation style and overall attitude towards them.

By using this approach it was possible to improve the user experience and usability methodology used, producing more reliable results and providing better user experience in the final version of the product as well as providing a pleasant experience during the testing of the product.

**Acknowledgments** We would like to kindly thank all the participants in the evaluation study for their participation and valuable feedback given.

## References

1. Albert, William, Thomas, Tullis (2013) Measuring the user experience: collecting, analyzing, and presenting usability metrics. Elsevier
2. Bachmayer S, Lugmayr A, Kotsis G (2010) Convergence of collaborative web approaches and interactive TV program formats. Int J Web Inf Syst 6(1):74–94
3. Brooke J (2013) SUS: a retrospective. J Usability Stud 8(2):29–40
4. Chorianopoulos K (2008) User interface design principles for interactive television applications. Int J Hum Comput Interact 24(6):556–573
5. ETSI (2012) Hybrid Broadcast Broadband TV, ETSI TS 102 796 v1.2.1 (2012–11)

6. ETSI Digital Video Broadcasting (DVB) (2010) Signalling and carriage of interactive applications and services in Hybrid broadcasting/broadband environments, ETSI TS 102 809 V1.1.1 (2010–01)
7. Guna J, Stojmenova E, Hauswirth D, Winkler R, Ninaus M, Pogačnik M (2013) WEB on TV: designing the user experience. V: STOJMENOVA, Emilija (ur.), LUGMAYR, Artur (ur.). Proceedings of the 1st Workshop on Defining a European Research Agenda on Information Systems and Management in eMedia Industries, Bled, Slovenia, 9th-13th June 2013, (International series on information systems and management in creative emedia, ISSN 2341–5576, no. 2013/1). Tampere: IugYmedia, str. 5–7.
8. Guna J, Kos A, Pogačnik M (2014) Gaining methodological insights from the user experience and usability study of an interactive HbbTV application. Proc. of the 7th Intern. Workshop on Semantic Ambient Media Experiences (SAME 2014) – Ambient Media Usability, Interaction, and Smart Media Technologies, In conjunction with NORDCHI 2014, Helsinki, Finland, 26th Oct
9. Hart, Sandra G (2006) "NASA-task load index (NASA-TLX); 20 years later." Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Vol. 50. No. 9. Sage Publications
10. Kalli S, Lugmayr A, Niiranen S (2004) Digital interactive TV and Metadata: future broadcast multimedia. Springer-Verlag, New York, doi:10.1007/978-1-4757-3953-4
11. Lowdermilk T (2013) User-centered design: a Developer's guide to building user-friendly applications. O'Reilly Media, Inc., Newton
12. Lugmayr A, Serral E, Scherp A, Pogorelc B, Mustaquim M (2014) Ambient media today and tomorrow. Multimed Tools App 71(1):7–37
13. Pauzié A (2008) "Evaluating driver mental workload using the driving activity load index (DALI)." Proc. of European Conference on Human Interface Design for Intelligent Transport Systems
14. SEE TV-WEB, (2015) Project brochure http://www.see-tvweb.eu/communication-material (lastaccessed on 18.2.2015)
15. Stojmenova E, Guna J, Hauswirth D, Winkler R, Ninaus M, Pogačnik M (2013) WEB on TV: cross-national user study for designing better user experience. V: LUGMAYR, Artur (ur.). Making sense of covering media. New York: Association for Computing Machinery, cop. 2007, str. 17–19
16. the TV-WEB project - Tackling the"(2015) Digital Divide" in SEE, http://www.see-tvweb.eu (last accessed on 18.2.2015)

**Dr. Jože Guna** graduated, defended his master's thesis and his Ph. D. thesis, all at the Faculty of Electrical Engineering, University of Ljubljana. His area of research focuses on Internet technologies, multimedia technologies and IPTV systems with special emphasis on user centred design, innovative user interaction modalities and designing the user experience, including gamification and flow aspects. Currently he is involved in a number of projects focusing on the development of intuitive user interfaces for elderly users of eHealth application, applications and games for the children and interactive multimedia HBBTV iTV applications. He is an expert in Internet, ICT and IPTV technologies and holds several industrial certificates from CISCO, Comptia and Apple, including trainer licenses from Cisco and Apple. He is a senior member and an official student counselor of the IEEE organization.

**Dr. Emilija Stojmenova-Duh** is a researcher at the Laboratory for Telecommunications within the Faculty of electrical engineering, University of Ljubljana, where she is involved in various research and development projects. She received her Ph.D. in Electrical engineering in 2013 for her doctoral thesis "User-centered Design for Multi-screen e-HealthApplications for Elderly People". One of the major dissertation contributions was the proposal for improving an existing standard (ISO9241- 210:2010), demonstrating the practical application of the results in an industrial standard. Her research work focuses mainly in the fields of user centred design and methodologies for evaluating user experience and usability for specific groups of users, such as: elderly people, children and people with disabilities. Previously, she was employed at Iskratel, Ltd. as a user experience manager, where she was responsible for the overall user experience in the company. She is "World Usability Day Slovenia" conference chair, IEEE Women in engineering (WIE) Slovenia section chair and an active member of IEEE, ACM UxPA, IxDA. Since September 2013, she is actively involved in the Demola Network as the head of RAZ:UM, Demola Slovenia operator. In Demola project ideas and needs come from the project partners, companies and organizations or international Demola Network partners.

**Dr. Matevž Pogačnik** graduated in 1997 and defended his PhD thesis in 2004 in the field of telecommunication and informatics at the University of Ljubljana. He is very active in the field of interactive multimedia services development, mores specifically in the fields of mobile applications, IPTV/smartTV applications and digital TV services. His recent research work has been mostly focused on improvement of user experience and introduction of new interaction modalities. He is also a member of the international organisation IEEE.