Iterative User Experience Evaluation of a User Interface for the Unification of TV Contents

Pedro Almeida$^{(\boxtimes)}$ 🆔, Jorge Abreu$^{(\boxtimes)}$ 🆔, Telmo Silva$^{(\boxtimes)}$ 🆔,
Enrickson Varsori 🆔, Eliza Oliveira 🆔, Ana Velhinho 🆔,
Sílvia Fernandes 🆔, Rafael Guedes 🆔, and Diogo Oliveira 🆔

CIC.DIGITAL/Digimedia, Department of Communication and Arts,
University of Aveiro, Aveiro, Portugal
{almeida, jfa, tsilva, varsori, elizaoliveira,
ana.velhinho, silvia.fernandes, rafaelguedes,
diogo.guerreiro}@ua.pt

Abstract. One of the key components of the design and development process of a User Interface (UI) is the User Experience (UX) assessment. It guarantees that the product meets the users' needs and provides the best interaction, while achieving usability goals and emotional reactions of motivation, arousal and dominance. In this scope, this paper describes the results driven from a User-Centered Design (UCD) methodology adopted in the development and evaluation of consecutive versions of semi-functional prototypes (of the UltraTV concept), with iterations that comprised an expert review and subsequent testing with users in a laboratory environment. This evaluation approach aims to achieve a proof of concept for the profile-based unification of traditional TV and Over-the-top contents offered in the same user interface. As an R&D project based on a partnership between the research field and the industry that aims to bring innovation to the interactive TV (iTV) domain, the UltraTV concept targets for the synergy between users and market interests. Despite the challenges of introducing an interface that unifies linear and nonlinear content in the same UI assuring the continuity of the UX contrary to the current app-based trend, the combined results from the experts' review and the inLab tests demonstrate the relevance and desirability of the concept as a potential solution for the future of iTV. The presented results provide valuable insights for further stages of field trials with end users, as well as to prove the feasibility and user demand for the profile-based unification for the next TV generation.

Keywords: Interactive television · Prototype evaluation · User interface
User experience · Content unification

1 Introduction

The development of a system corresponds to an iteration process, as it considers opinions from different evaluators, in distinct phases of the prototypes, for identifying issues to be improved and enhancing the global user experience [1].

Additionally, it is crucial to the success of a system to timely test it, since its development benefits substantially from the gathered insights. These tests also offer the

opportunity to evaluate design alternatives in low fidelity prototypes, allowing the iterative refinement of different versions presented to the evaluators [2].

A review by experts plays an important role in this process, as it provides access to the evaluators' professional experience (simultaneously as developers, researchers and users), representing a value-added assessment at both levels: in the development process and in a final solution. Additionally, laboratory tests can simulate a domestic environment, providing an experience resembling a real living room, contributing for gathering suggestions to make improvements in the system and guarantee the production of a high-fidelity prototype that meets the users' needs. Therefore, a variety of tests should be made throughout the development process [2].

The referred iterative development process has a special relevance in the ongoing academic and industry initiatives that aim to streamline the future of the TV ecosystem. The current television scenario is going through fundamental changes in the way viewers get access to TV contents, increasingly supported by Video On Demand (VoD) services in addition to the traditional lineup of TV contents. This transformation on the viewing habits is supported by recent solutions, granting access to the available contents anytime and anywhere [3, 4]. These solutions include the use of mobile devices, working as secondary screens [3, 5], and many times as the primary displays for watching TV. Also, the services of traditional Pay-TV operators have been trying to adjust to the contemporary scenario. To grant its clients with access to linear contents at alternative schedules, Pay-TV operators allow pre-programed TV to become more flexible by introducing services like "catch-up TV" and "time-shift" [6, 7]. This framework leads to enormous changes in the TV offer, increasing the challenges of commercial players that are trying to cope with new users' expectations by offering more engaging systems supported on new paradigms of User Interfaces (UI). In this scope, the consortium of the UltraTV project (involving Altice Labs as an IPTV player, University of Aveiro and Instituto de Telecomunicações - a R&D institution in the field of telecommunications) is designing an iTV concept, with the most advanced features on the market, capable of serving as a basis for a new generation of a Pay-TV service.

In this paper, a systematization of the evaluation tests performed to consecutive versions of the UltraTV prototype is presented, addressing two types of tests: a review by five experts, organized in two panels (one in Chicago and the other in Aveiro); and an evaluation in a laboratory environment with 20 participants (which took place during a week in Aveiro). It should be noted that the version of the prototype tested in the lab with users integrated some improvements, resulting from the suggestions of the experts, while other aspects were maintained to test the opinion of the users in order to confirm some of the prior insights. In this sense, two versions of the prototype, designated by Version 1 (tested with experts) and Version 2 (tested inLab with users) are described in this article.

The paper is structured as follows. The next Sect. (2) presents the state of the art concerning current viewing habits, iTV industry trends and User Experience (UX) assessments in the iTV context. In the next section, the UltraTV project is presented based on what was named the "TV Flow", as a foundational concept for a profile-based unification approach translated into a user interface to access content from different sources. Section 3 presents the evaluation goals and methodology used for the validation of the prototypes. In Sect. 4 the results are discussed, namely the assessment

made by the experts and the outputs from the inLab tests. Finally, the last Sect. (5) highlights the most relevant considerations and suggestions for future work.

2 State of the Art

2.1 Viewing Habits and Industry Trends

The television scenario has been suffering significant changes, regarding the transformation on viewers' habits. To cope with this new context, Pay-TV operators provide new features to increase enjoyment and to personalize the users' TV experience, such as catch-up TV, time shifted, and Video-on-Demand (VoD). This context is sustained by the growing availability of audiovisual technology [7], which makes it possible to watch any content, anywhere, at any time [4]. From the several UI design approaches leveraged by the iTV industry that try to adjust to the current scenario and provide an engaging experience there are two that are worth being mentioned. Supported by a flexible grid, the Argentine operator Cablevisión[1] proposes a different layout approach with the introduction of the Flow solution, in an all intuitive UI, combining linear content, OTT, Time-shift TV services with content filtering according to a user's preferences. In turn, Android TV[2] proposes the unification of contents from Google sources and for recommendations on the home screen, using a carrousel menu with different contents in a same line and allowing the connection to other Google features.

Moreover, innovative layouts that emphasize the dynamics of the navigation supported by transition effects have been presented, such as the three-dimensional effect in Mur Vidéo[3] by Voo, the diagonal carrousel navigation of the smart TV LG webOS[4], the horizontal navigation with circular items, of Frog[5] interface by Wyplay, or the disruptive interfaces by Cisco, like the Infinite Video Platform[6], based on video masks. All this scenario leads to the hybridization of the ways TV is viewed and provide changes on the paradigm of graphical approaches (centered on layout, typographic treatment, animations, transitions and effects) associated with navigation and visual feedback given to the user, leading to an easy and intuitive but also captivating and immersive UX [8].

2.2 Evaluations in ITV Context

It is possible to relate UX to several different contexts, leading to a nonconsensual concept in what concerns the meaning of the term [9, 10]. Therefore, many are the definitions related to UX, which are strictly related to the issues correlated to the respective individual area. Following this perspective, Law defines UX as follows:

---

[1] Cablevisión Flow: https://www.youtube.com/watch?v=IE0n-WlGHYg.

[2] Android TV: https://www.android.com/tv/.

[3] Voo Mur Vidéo - https://www.youtube.com/watch?v=dwLK2stDH0g.

[4] LG webOS 3.5 - https://www.youtube.com/watch?v=v5471kC8heE.

[5] Wyplay Frog - https://www.youtube.com/watch?v=cNSkoeDaGDY.

[6] Cisco Infinite Video Platform - https://www.youtube.com/watch?v=5LrVIablSJk&t=280s.

"User Experience (UX) is a catchy as well as tricky research topic, given its broad applications in a diversity of interactive systems and its deep root in various conceptual frameworks, for instance, psychological theories of emotion" [10]. Furthermore, Bernhaupt also highlights that UX in the iTV domain is related to four dimensions - aesthetics, emotion, stimulation and identification - measurable by a set of instruments, to assess non-instrumental and emotional reactions [11]. Also, ISO designates UX as: "A person's perceptions and responses that result from the use and/or anticipated use of a product, system or service" [12]. However, due to several possible ramifications, this definition originates more specific interpretations, such as: "User Experience explores how a person feels about using a product, i.e., the experiential, affective, meaningful and valuable aspects of product use" [13].

Regarding the UX process, the evaluation of a prototype by experts represents a crucial step towards the assessment of instrumental qualities related with a product's usability and a quick mean to gather relevant improvement suggestions. In this regard, two facts are important to highlight: (i) the evaluation of the UI should be as early as possible in order to offer designers the chance of getting feedback for the redesign process [14]; (ii) the selection of suitable evaluators considering their background, expertise and previous experience with similar systems [15]. After the definition of the experts, it is important to define methods and the tasks that will be applied [16].

Concerning the evaluation process with users on the TV ecosystem scenario, there are specificities that need to be considered before the realization of tests. The subjectivity of the user is a crucial factor to considerate, encompassing temporal, spatial, social, personal and technological factors, as well as the literacy level of each user. Simultaneously, many are the components that influence the UX in the iTV context, such as the Set-Top Box (STB) performance, the remote control, second-screen devices (smartphone and tablet), the television itself, among others contextual factors. Therefore, in the evaluation planning, is crucial to encompass usability questions, together with emotional factors as important topics to be assessed. Hence, it is necessary to use tools in order to comprise, not only instrumental qualities but others important aspects regarding the non-instrumental qualities of the UX [9, 17].

Among the several methods for usability evaluation, some of the most common are the heuristic evaluation, cognitive walkthrough and guideline review [16]. The think aloud protocol is also frequently used to encourage participants to express their opinions during the real-time experience. The heuristics consider a set of UI related principles. A cognitive walkthrough methodology is similar to a heuristic evaluation but with an emphasis on a task scenario that evaluators need to perform in the UI. The participants must go through predefined steps and identify issues concerning the interface and navigation. Finally, a guideline review involves having an evaluator comparing an interface against a detailed set of guidelines.

The experts' evaluation in the UltraTV project considered tasks converted to scenarios and contexts with specific navigational paths in order to create a guided interview for the evaluators, using a storyboard in a semi-functional prototype created on the Marvel platform accessed through an application provided by Apple TV. Additionally, in both, laboratory and experts' evaluations, the protocol comprised cognitive walkthrough and the think aloud methods.

On the second stage of the UltraTV assessment, the laboratory tests were focused on a triangulation of instruments [17] to collect feedback about the semi-functional prototype. For the perception of instrumental qualities (e.g. controllability, effectiveness, learnability), the SUS-System Usability Scale [18, 19] and the pragmatic dimension of the AttrakDiff were used. Regarding the evaluation of non-instrumental dimension, the hedonic dimension of AttrakDiff[7] (e.g. aesthetics, identification) was used; and, finally, to achieve emotional information reactions (e.g. satisfaction, motivation and control), the SAM-Self-Assessment Manikin [20] was the validated scale chosen to be applied in laboratory tests.

3   UltraTV Project: Prototype Development and Evaluation Methodology

The UltraTV project adopted an iterative and participatory design approach with the aim of developing a TV ecosystem that meets the contemporary demands regarding viewing habits and industry trends. The main goal is to consolidate an interactive TV ecosystem that supports the most innovative services while integrating audiovisual content (from different sources) in a unified way. In this sense, this project is focused on facilitating the access to On Demand content in an integrated UI, surpassing the linear and traditional organization of the channel's line-up. For this, it pursues the iterative design of user interfaces, as well as the validation of its feasibility through several testing methods (experts reviews, inLab tests and field trials). These goals have been translated into an UI proposal named TV Flow. The TV Flow aims to promote a consistent and fluid UX while interacting with an engaging interface, providing access to profile-based recommended content from different content sources. Additionally, the TV flow concept can be considered a different approach to zapping by promoting discovery of new contents, namely of content from players that the clients typically don't use, or, inversely, by promoting an easy access to frequently visited content sources, avoiding changing between different apps. With this aggregation and unification in mind, the UI intends to reconfigure the way people perceive and use the TV (see Table 2). Through a unified UI, content from different sources, namely from YouTube, Netflix, Facebook Videos along with traditional linear TV content is provided to the user. The content is laid out in a grid categorized in columns, according to genres (e.g. Sports, Information, Movies and Series) and sources (e.g. Live TV, YouTube, Facebook Videos, Netflix). The proposal of a disruptive UI for the unification of TV content aims precisely to enhance the entertainment experience offered to different consumer profiles – on the one hand by providing content beyond the traditional broadcast channels to the classic TV consumer (promoting the diversification of his choices), and on the other hand by aiming to capture the attention in TV content and improve the quality of OTT content viewing on the large TV screen for those who are used to watch videos using computer and mobile displays.

---

[7] AttrakDiff: http://attrakdiff.de/.

The initial tasks of the preliminary phase consisted in defining the requirements and guidelines of the intended solution. After the first sketches, low-fidelity mockups were created, tested and discussed internally by the different team members in brainstorming sessions. A medium fidelity prototype was then developed in the Sketch software, being later transposed to the Marvel application for the experts' review (see Table 2 – Version 1). The design of this first prototype was based on an exploratory interface supported on a grid-based navigation using the four remote control buttons and was then tested with the command belonging to the Apple TV (using the touch surface instead of the directional keys). As for as the prototype version tested inLab, it was developed using Luna, a JavaScript framework, installed in an Android box connected to the TV with a regular remote-control interaction (see Table 2 – Version 2).

The evaluation of the two versions of the prototype (Experts review for Version 1 and inLab tests for Version 2) were crucial to a better understanding of the system, namely to find how the prototype could be improved before its final assessment in a Field Trial.

3.1 Experts Review

In the UX evaluation process of preliminary prototypes, the expert's review represents a crucial step in the evaluation of a product and an agile method to gather relevant suggestions. For this to happen effectively two factors need to be taken into account: (i) UI assessment should be performed as early as possible to allow valid feedback for the redesign process; (ii) the selection of suitable evaluators considering their professional area and their experience with similar systems. Following the selection of experts, methods and tasks that effectively result in an overall analysis of the usability of the interface and the user experience should be provided.

For the UltraTV expert review [21], the analysis of users and tasks was converted to scenarios with specific navigational paths to create a cognitive walkthrough using a storyboard, allowing also free explorations whenever required. The session began with a presentation about the project and the evaluation objectives, complemented by a video that included transitions and animations of navigational menus, to give the evaluators a clear idea about the look & feel of the interface. During both the free explorations and the cognitive walkthrough of the prototype, the participants were encouraged to think aloud.

The UI assessment focused mainly on a qualitative approach to gather insights about navigation and graphical features to guide the redesign process. To complete this review efficiently, Nielsen's recommendation to use 5 experts was followed, since this number allows the identification of 85% of usability problems and was considered the most viable option for this qualitative test [22, 23].

3.2 Tests in Laboratory

Taking into consideration the recommendations obtained in the evaluation by experts and subsequent developments implemented in the system, the following objectives were defined for the tests with users in a laboratory environment (simulating a regular living room):

- determine the level of interest and acceptance of the overall proposal for unification of content from different sources based on user profiles;
- evaluate the intuition and consistency of the UI and navigation and grid-based organization;
- validate the keys of the remote control used to interact with the system;
- evaluate the overall appearance of the interface, including graphic aspects such as color, iconography, shapes and effects;
- determine the relevance of specific features (display modes, icons, profiles…);
- identify global and localized usability issues;
- gather information about users' needs and suggestions.

To achieve the established objectives, exploratory assessments (cognitive walkthrough and think aloud) and validated usability and UX scales (SUS, SAM and AttrakDiff) were applied. Table 1 shows a concise description of each phase of the inLab tests.

Table 1. Activities carried in the tests of the UltraTV prototype.

Phases	Activity	Description
Introduction	Characterization Questionnaire	Focused on demographics, TV consumption habits and devices usage
	Project introduction	Introduction about the project and the structure of the test session
	Free exploration	Free exploration of the application for a few minutes
Exploration	Cognitive walkthrough	14 tasks mainly focused on navigation and features usage
	Exploration with think aloud	Free exploration of the application for a few minutes
Assessment	SUS (System Usability Scale)	Global usability overview using a ten-item attitude Likert scale
	SAM (Self-Assessment Manikin)	Non-verbal pictorial assessment that measures the satisfaction, motivation, and control associated with a person's affective reaction
	AttrakDiff	Instrumental and non-instrumental issues evaluation
	Interview	Interview including open-ended and close-ended questions

3.3 Participants Characterization

Regarding the expert's review, five evaluators were selected. Two experts with an academic background in technology, namely in UX and UI, being professors at the Illinois Institute of Technology of the Institute of Design- Chicago/USA; three experts developers of IPTV and iTV systems in Altice Labs (Portugal), with experience in UX and Human-Computer Interaction (HCI).

Regarding the inLab tests, a total of 20 participants took part in the tests sessions (10 males and 10 females), aged between 20 and 53 years old (average age = 28.5).

The sample consisted of individuals with different qualifications: Primary education (1), Secondary education (7), Bachelor's degree (7), Post-graduation (1), Master´s degree (3) and PhD (1). Considering the professional occupation, it included: students (10); research fellows (5); employees (2): unemployed (1): retired (1): and freelancers (1). Regarding the TV consumption habits, 80% of the participants stated to have at least one television at home and to use a TV subscription service. Only 15% of participants referred not having access to any pay-TV service. Within the selected sample, the clear majority watched television at home (70%).

4   Results and Impacts on Prototype Development

For the presentation of the results the chronological order of the evaluations will be followed.

4.1   Results from the Experts' Review and Impacts on the Prototype

Considering the experts' evaluation, the analysis of the interface and corresponding navigation (visualization mode; menu and profile; organization and unified content) and of the look & feel (contextual menu|animations and transitions|) was carried by the 5 experts.

The experts agreed with some of the proposals shown, namely the concept of the grid, the need to focus on active content and the need to better emphasize the selected content. The experts also expressed positive feedback on profiling and contextual menus. However, some issues originated divergent opinions. An example was the relevance of the two viewing modes (Table 2 – a1 and b1), being mentioned as confusing by some experts and, contrarily, being valued by others as a feature of customization and discovery. Only one expert clearly rejected the wide view mode, stating that he would only use the zoom mode. Other experts said that they would switch between the two modes as they had a clear idea of what they would like to see using zoom mode (Table 2 – b1), or whether they were open to other possible options using wide mode (Table 2 – a1). In terms of the unification of content, the main functionality of this solution, it was approved by all.

Following the experts' comments, several adjustments to the prototype were defined and a new version (Version 2) was created as a synthesis of those comments. This was the version used in the following evaluation step, the inLab tests Table 2 presents the differences between the two versions of the prototype, comparing four main areas of the UI, namely the home screen grid in the wide visualization mode (Table 2 – A), the home screen grid in the zoom visualization mode (Table 2 – B), the full screen information menu at the bottom of the screen (Table 2 – C) and the contextual menu also in full screen (Table 2 – D). Issues concerning hierarchy and contrast were addressed in the home screen grid to reinforce the vertical reading of the columns.

The header from Version 1 (Table 2 – a1) was replaced in Version 2 (Table 2 – a2) by a blob shaped menu with a fluid and animated behavior when changing status. This menu allows shifting the view mode, the profile and to access the search feature. The labels that identify each category/column were given a different graphic treatment to

Table 2. Comparison of prototype versions (Version 1 and 2)

Prototype – Version 1 (Experts Review)	Prototype – Version 2 (inLab Tests)
(A) Home screen – Wide view	
(a1)	(a2)
(B) Home screen – Zoom view	
(b1)	(b2)
(C) Full screen – Bottom Contextual Menu	
(c1)	(c2)
(D) Full screen – Bottom Contextual Menu	
(d1)	(d2)

avoid being interpreted as a menu. Furthermore, fewer content cards were exhibited in the grid reinforcing the scale focus given to the selected card. The central column, that displays the active topic, was also emphasized using a more vivid color. The zoom view mode (b2) was simplified with a cleaner approach and the focus on the selected video was also improved being displayed in a larger size and with the category label overlaid (e.g. "Mix TV" label that refers to a group of suggestions from catchup TV). The contextual menus that provide additional information (C) and offer additional features to interact with the content (D) were also redesigned. The bottom contextual menu (c2) suffered minor changes mainly in color and opacity becoming less disruptive. Additionally, the functionalities menu (d2) was placed at the top of the screen, following the same blob shape and behavior of the home screen menu and including five options instead of four (restart, record, dislike, next content, more info). The experts mentioned that displaying the contextual menu at the centered with a darker texture over the content (d1) represented an unnecessary spotlight on this menu becoming a disruptive obstacle to the content viewing experience. Modifications to the UI, like the ones mentioned above, regarding graphic elements, textures, color, text styles, animations and navigation were extremely relevant to improve the overall UX and provide a distinctive look & feel to the system. Based on these results, a new UI (Table 2 – a2, b2, c2, d2) was adjusted to allow its integration into the subsequent prototype to be evaluated in laboratory.

4.2 Results from InLab Tests and Impacts on the Prototype

Regarding the inLab tests, in the Cognitive walkthrough session, the success rate (on 14 tasks) was classified between the parameters 1 "without problems" and 2 "with some difficulty", with the average of all participants being 1.5 (being 1 "without problems" and 5 "did not complete") which shows no significant problems with the use of the prototype.

Participants were then asked to answer to a triangulation of instruments for UX evaluation [11, 17], namely the scales SUS, AttrakDiff and SAM, whose overall results are shown in Table 3:

In terms of instrumental qualities, the prototype scored 83,63 in the SUS scale (ranging from 0 to 100) and 1.64 in the Pragmatic Dimension of AttrakDiff (ranging from −3 to 3). This score reflects the user´s controllability, effectiveness and

Table 3. Application score according to the SUS, SAM and AttrakDiff instruments.

Instrumental qualities	Non-instrumental qualities			Emotional reactions			
SUS (0 to 100)	AttrakDiff (−3 to 3)			SAM (1 to 9)			AttrakDiff (−3 to 3)
	PQ	HQ-S	HQ-I	Sat.	Mot.	S. of C.	ATT
83.63	1.64	1.24	1.49	7.8	7.5	7.4	2.01
UX Dimensions	Stimulation	Identification		Emotion			Aesthetics

learnability regarding the system, revealing that in this test stage the prototype was already getting favorable results.

According to the SUS classification scale, see Fig. 1, the average value was 83.63 ("Good"), although very close to the classification considered "Excellent" (85.5), which confirms a high level in terms of system usability, since the minimum acceptable value starts in the second quartile, which corresponds to a score of 62.6.

Fig. 1. Diagram with the prototype final score of SUS.

Regarding Self-Assessment Manikin (SAM) results, consolidating what was revealed in the SUS questionnaire, the participants demonstrated a positive emotional reaction in what concerns the use of the UltraTV prototype. The best scores were given to the satisfaction parameter, with an average of 7.80 in 9, and the motivation and control topics with nearly the same average, 7.55 and 7.45 respectively. Although in general good results was obtained, the need for improvements in the feeling of control over the application was detected, since the items related to this issue had lower scores. Similarly, it is necessary to remember that the tested prototype was not fully functional (although the grid interface was functional, users were shown areas with no possible interaction, which may have influenced the control quality).

Regarding the results obtained from the AttrakDiff scale, the average values of the 4 dimensions were calculated with all dimensions getting high scores (Fig. 2). Considering the instrumental aspects, the PQ obtained good values, following the usability results achieved by the SUS. However, the best results came from the Attractiveness (ATT) dimension, which is a non-instrumental quality and is strictly related to the aesthetic issues of the prototype. With an average value of 2.01, the value of the ATT dimension matches the qualitative feedback on the UI look & feel gathered in the interviews. In this same dimension, the results of the pairs of words "bad-good", "rejecting-inviting", and "disagreeable-likeable" were highlighted.

For non-instrumental dimensions, the HQ-I factor indicates the extent to which the system allows the user to identify with it, while the HQ-S dimension indicates originality, interest and encouragement towards the system. The HQ prototype scores in these dimensions (1.24 in HQ-S and 1.49 in HQ-I) show high rates of hedonic results, meaning
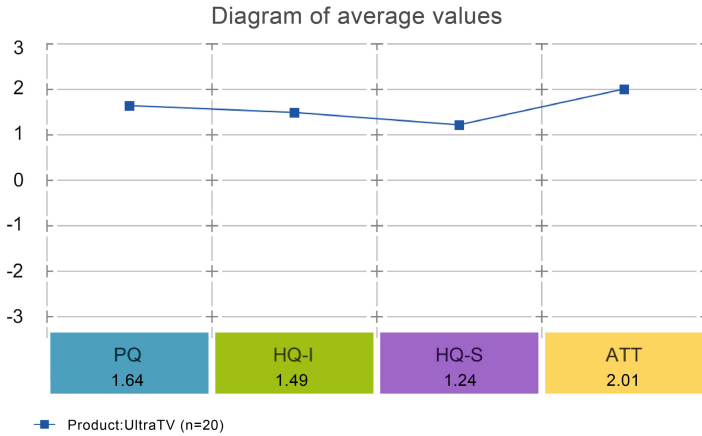
Diagram of average values

	PQ	HQ-I	HQ-S	ATT
	1.64	1.49	1.24	2.01

Product:UltraTV (n=20)

Fig. 2. The average values of the 4 AttrackDiff dimensions.

that the user identifies with the product, feels motivated and stimulated, and considers the product desirable (Fig. 3). In the HQ-I dimension, the "unpresentable - presentable" pair obtained the best result, while in the dimension HQ-S the pair "undemanding - challenging" was best classified. The AttrakDiff results, along with the scores obtained from

Portfolio-presentation

www.attrakdiff.de

hedonic quality (HQ)	too self-oriented	self-oriented	desired
		neutral	task-oriented
	superfluous		too task-oriented

pragmatic quality (PQ)

www.attrakdiff.de

Product:UltraTV (n=20)
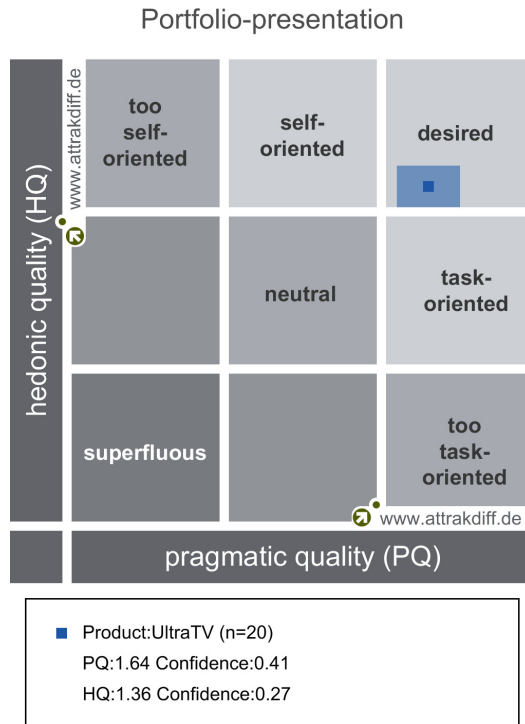PQ:1.64 Confidence:0.41
HQ:1.36 Confidence:0.27

Fig. 3. Portfolio-presentation with confidence rectangles.

the SAM questionnaire, reveal that, although the participants consider the prototype undemanding, there are still problems in the manipulation of the system that can be subject to a more in-depth analysis using the users' feedback provided in the interviews.

Regarding this UI proposal, according to the inLab test evaluators the Version 2 of the prototype (Table 2 – a2, b2, c2, d2) was considered very thorough, intuitive and desirable, without being too complex. However, further recommendations were drawn from the inLab tests to be considered for an improved version to be used in the field trial (Version 3). Concerning the content unification, the need for the content to be hierarchized according to usage and consumption habits, placing at the top of the grid the more relevant content to the user, was identified. The results also showed the need for the search tool to be unified in a way that searches are global and related to all sources of content. Tutorials should be provided to elucidate about the structure and organization of the system. The need for a filtering system to simplify and customize the content presentation was also detected. Finally, concerning the customization of the system, the inclusion of a social component, like friends' recommendations was also mentioned.

5   Conclusions and Future Work

The preliminary evaluation made by the experts and later in the inLab tests allowed the team to gather a solid and highly relevant set of results crucial for the improvement and production of a functional prototype. The different results presented are contributing to the newer versions being developed. Subsequent to these evaluation stages, a proof of concept was achieved and iterations of the prototype are consolidating an improved version to be tested in the homes of potential users. Through this user centered design methodology, the UltraTV project aims to create a complete viewing ecosystem for television and video consumers. It is our believe that such conceptual development has the potential to make an important contribution to change the current television consumption paradigm. With the current and future evolutions, unified content and easy access to OTT sources, along with online social media content, we hope that the UltraTV project will have a role in the debate on what the future of the television system will be like at both the academic and the industry level.

References

1. Norman, D.A.: The Design of Everyday Things. Doubleday, New York (1990)
2. Interaction Design Foundation: The Basics of User Experience Design (2002)
3. Abreu, J., Almeida, P., Teles, B.: TV discovery & enjoy: a new approach to help users finding the right TV program to watch. In: Proceedings of the 2014 ACM International Conference on Interactive experiences for TV and online video, pp. 63–70. ACM (2014)

4. Vanattenhoven, J., Geerts, D.: Broadcast, video-on-demand, and other ways to watch television content: a household perspective. In: Proceedings of the 2015 ACM International Conference on Interactive Experiences for TV and Online Video, pp. 1–10. ACM (2015)
5. Vanattenhoven, J., Geerts, D.: Designing TV recommender interfaces for specific viewing experiences. In: Proceedings of the 2015 ACM International Conference on Interactive Experiences for TV and Online Video, pp. 185–190. ACM (2015)
6. Gonçalves, D., Costa, M., Couto, F.M.: A large-scale characterization of user behavior in cable TV. In: 3rd Workshop on Recommendation Systems for Television and Online Video (2016)
7. Abreu, J., Nogueira, J., Becker, V., Cardoso, B.: Survey of Catch-Up TV and other time-shift services: a comprehensive analysis and taxonomy of linear and nonlinear television. Telecommun. Syst. **64**, 57–74 (2017)
8. Cardoso, B., de Carvalho, F.: Trends in TV: OTT, Apps, Ads. Homepage http://www.alticelabs.com/content/07_TV.pdf. Accessed 22 Nov 2017
9. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behav. Inf. Technol. **25**, 91–97 (2006)
10. Law, E.L.-C.: The measurability and predictability of user experience. In: Proceedings of the 3rd ACM SIGCHI Symposium on Engineering Interactive Computing Systems - EICS 2011 (2011)
11. Bernhaupt, R., Pirker, M.: Evaluating user experience for interactive television: towards the development of a domain-specific user experience questionnaire. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8118, pp. 642–659. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40480-1_45
12. ISO 9241-210, Ergonomics of Human–System Interaction – Part 210: Humancentred Design for Interactive Systems (2010)
13. User Experience (n.d.). https://en.wikipedia.org/wiki/User_experience. Accessed 22 Nov 2017
14. Demetriadis, S., Karoulis, A., Pombortsis, A.: "Graphical" Jogthrough: expert based methodology for user interface evaluation, applied in the case of an educational simulation interface. Comput. Educ. **32**, 285–299 (1999)
15. Jin, J., Geng, Q., Zhao, Q., Zhang, L.: Integrating the trend of research interest for reviewer assignment. In: International WWW Conference, pp. 1233–1241 (2017)
16. Rohrer, C.P., Boyle, F., Wendt, J., Sauro, J., Cole, S.: Practical usability ratings by experts: a pragmatic approach for scoring product usability. In: CHI 2016, vol. 38, pp. 333–358 (2016)
17. Abreu, J., Almeida, P., Silva, T.: A UX evaluation approach for second-screen applications. In: Abásolo, M., Perales, F., Bibiloni, A. (eds.) Applications and Usability of Interactive TV, vol. 605, pp. 105–120. Springer, Cham (2016)
18. Brooke, J.: SUS-A quick and dirty usability scale. In: Jordan, P.W., Weerdmeester, P.W., Thomas, P.W., McLelland, I.L. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor and Francis, London (2015)
19. Martins, A., Rosa, A., Queirós, A., Silva, A., Rocha, N.P.: European Portuguese validation of the usability scale (SUS). In: 6th International Conference on Software Development and Technologies for Enhancing, pp. 293–300 (1996, 2015)
20. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exper. Psychiatry **25**(1), 49–59 (1994)
21. Almeida, P., Abreu, J., Oliveira, E., Velhinho, A.: Proceedings of 6th Iberoamerican Conference on Applications and Usability for Interactive TV, pp. 59–70. University of Aveiro, Aveiro, Portugal (2017)
22. Nielsen, J.: Why You Only Need to Test with 5 Users (2000). https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/. Accessed 22 Nov 2017
23. Nielsen, J.: How Many Test Users in a Usability Study? https://www.nngroup.com/articles/how-many-test-users/. Accessed 22 Nov 2012