Field Trial of a New iTV Approach: The Potential of Its UX Among Younger Audiences

Ana Velhinho(✉), Sílvia Fernandes, Jorge Abreu, Pedro Almeida, and Telmo Silva

Digimedia, Department of Communication and Arts, University of Aveiro, Aveiro, Portugal
{ana.velhinho, silvia.fernandes, jfa, almeida, tsilva}@ua.pt

Abstract. The actual video consumption behaviors are blurring the traditional boundaries between linear and non-linear viewing, leveraged by On-demand and Over-the-Top (OTT) content offers. The Managed-Operated Networks (MON) adapted to this shift by providing flexibility to content access via Catch-up and Time-shift TV services. In addition, manufacturers of Smart TVs and high-end media centers (e.g. Apple TV) are offering their commercial solutions with a silo-based approach, lacking a fluid User eXperience (UX) when the user switches between proprietary applications (apps) displaying different settings and User Interfaces (UIs). In this framework, OTT providers have been taking the lead on setting the trends with partially unified cross-source UIs, enriched with personalization features to enhance the UX. These trends also contrast with Pay-TV solutions based in UIs mostly oriented to the traditional TV channel's line-up. Based on this opportunity, an operator-based iTV solution, delivered over a set-top box (STB), was developed within the UltraTV R&D project. This initiative brought together the academic and the industry fields to design, develop and validate a concept of a profile-based and cross-source recommendation UI (offering, at the same level, content from linear and non-linear TV, Netflix, YouTube and Facebook videos). After an iterative evaluation through expert reviews and laboratory tests, a field trial validation, with end users, of a fully functional prototype was set. Qualitative data collection methods along with a triangulation of UX scales were applied to gather opinions and measure usability, hedonic and emotional parameters. The results were arranged by age groups and viewing dynamics to reveal motivational indicators related to the achieved viewing experience. The unified access to different sources fostered content discovery and was particularly valued by younger audiences. The outcomes from this empirical study aim to provide valuable contributions to push the next generation of television platforms and bring younger audiences back to the TV screen.

Keywords: Content unification · Field trial · Interactive television · User experience · Viewing behaviors

M. J. Abásolo et al. (Eds.): jAUTI 2018, CCIS 1004, pp. 131–147, 2019.
https://doi.org/10.1007/978-3-030-23862-9_10

1 Introduction

The growing relevance of non-linear video over linear TV is highlighted as the main transformation within audiovisual consumption in recent studies [1], primarily due to the increase of OTT offers and the convenience of mobile viewing as a complement/alternative to the TV screen. This framework has promoted new behavioral phenomena like a) the "cord-cutters" (users that drop their Pay-TV services in favor of streamed OTT content making the regular TV service disposable) and b) the "binge-watching" (refereeing to a continuous marathon viewing of several episodes from the same show). Both phenomena are mainly associated to the users' groups "Mobility Centrics" (that essentially use mobile screens) and "Screen Shifters" (that consume all kind of TV and video content across multiple devices and anywhere) which are pushing new TV and video services [1]. To cope with this new context where OTT players are gaining space, Pay-TV suppliers react with commodity bundle TV and Internet services which include Video-on-Demand (VOD), time-shifted, and catch-up features also available on the move, anytime and anywhere [2–4].

Often this demanding new type of viewers is being studied from the perspective of generational groups [5], which are also associated with viewing behaviors classified by Ericsson Consumer Lab reports [1] TV groups. The "Silent Generation" (ages 65+), "Baby Boomers" (ages 50–64)" and "Generation X" (ages 35–49) are the ones that preferably watch video on the TV screen, mostly sports, news and movies [1]. Contrariwise, computer, tablet and mobile phone-based viewing is more popular among youngest consumers – "Millennials" (ages 21–34) and "Generation Z" (ages 15–20) – not because they totally dismiss the TV screen but because they adopt multitasking practices. This also has an impact on the number of hours dedicated to video consumption, being a total of 33 h per week attributed to groups of users between 16–19 years old, with 54% of this time dedicated to On-demand content. Users between 20–24 years old have similar behaviors. In contrast, the group of users between 45–49 dedicate only 31% of their time watching On-demand content (from a total of 29 h per week) whereas the 60–69 age group have a similar total amount of viewing time, however with only 21% of this period dedicated to VOD [1].

Within this framework, the personalization of the experience, along with a content-first approach delivered through a UI based on aggregation and unification of content and sources may represent leading features to meet current users' media behaviors. Such topics are targeted by an R&D project that brings together the academic and the industry fields to design, develop and validate a concept of a profile-based and cross-source recommendation UI that offers, through a STB based prototype, TV content at the same level with videos from Netflix, YouTube and Facebook [6, 7].

The main goal of this paper is to present insights provided by younger generations towards a unified iTV solution based on the results obtained through an empirical study with the UltraTV prototype. The prototype was developed according to an iterative User-Centered-Design (UCD) methodology [8] that included three evaluation phases: (i) an expert's review [9, 10], (ii) tests in laboratory [6] and (iii) a Field Trial (FT) [11].

After integrating the feedback from previous tests, the focus of the FT procedure was to validate the unification concept and assess the overall User eXperience (UX) of the system using a triangulation of validated instruments (described in section three): the System Usability Scale (SUS); the Self-Assessment Manikin (SAM), and; the AttrakDiff. Also, a characterization questionnaire was applied. Complementary to these instruments was the permanent monitoring of the interactions during the FT and phone interviews conducted afterwards.

The paper is structured into five sections. After this introductory part, the next section presents recent content unification solutions in the iTV domain, including the UltraTV R&D project. In the following section, the FT of the UltraTV prototype is described, detailing the UI of the tested version, the tests' procedures and the sample characterization. The final sections present and discuss the main results regarding the younger audience's behaviors, culminating with the conclusion that summarizes the contributions of the study.

2 Content Unification in the iTV Domain

Recently, online players have developed niche business models relying on OTT streaming technology. To face these OTT newcomers, cable Pay-TV subscriptions have started to offer additional On-Demand services. Despite the fast growth of such players, limitations on getting traditional TV channels and live content [12] kept them from taking over the TV market. Those restrictions prevent a seamless UX, perpetuating a behavior of using multiple apps displaying different settings and UIs. The lack of an integrated search feature and cross-source recommendations is also a limitation that leads to users having a scattered audiovisual consumption across devices. The interdependent dynamics between channels and Pay-TV operators and OTT providers is bringing both sides closer together in favor of more balanced solutions regarding a personalized and content-first viewing experience.

However, it was not before the turn to 2018 that new industry releases seemed to acknowledge the benefits from the synergies between OTT platforms and cable companies to more suitable services to meet the viewers' demands. The SkyQ[1] solution, from the UK Sky operator, led a paradigmatic change towards a cross-content approach by proposing a UI allowing the user to switch between live TV and Netflix content seamlessly (see Fig. 1). The other way around was the strategy adopted by the VOD platform Hulu[2], that integrated live content and channels alongside their original on-demand content, offering a "Lineup" section to display personalized content for different profiles (see Fig. 1).

[1] Sky Group: http://bit.ly/2LcigCg.

[2] Hulu Homepage: https://www.hulu.com/live-tv.

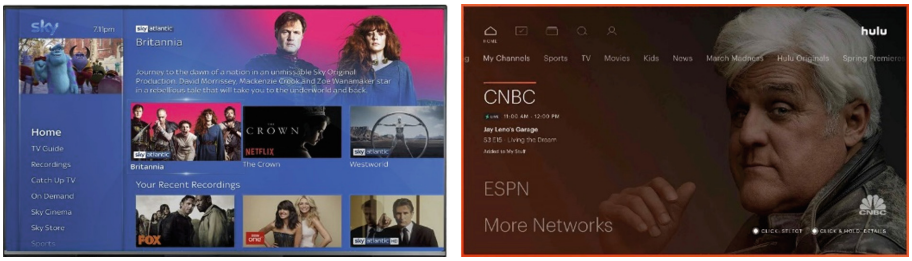

Fig. 1. Sky Q UI (left) and Hulu UI (right)

From beginning of 2017 until July 2018, an R&D project (UltraTV) unfolded from a partnership between the industry (a leading IPTV operator) and the academy. The goal was to target the new TV generation focused on the unification and personalization, by combining in the same UI, TV and OTT content (see Figs. 2, 3 and 4). With a MON-based framework [3] supported by the operator's offer, it provides unified access to a diversified bundle of content sorted by genres and sources. Content is displayed on a card-based grid with axis navigation and a fluid blob menu redirecting to the system's main features: "My Content" area (favorites and keep watching), filtering, profiles, unified search, and settings (see Fig. 3). Cross-content recommendations are also offered to foster the discovery of content (Fig. 4). Experts and end-users validated this solution during several UX assessment stages.

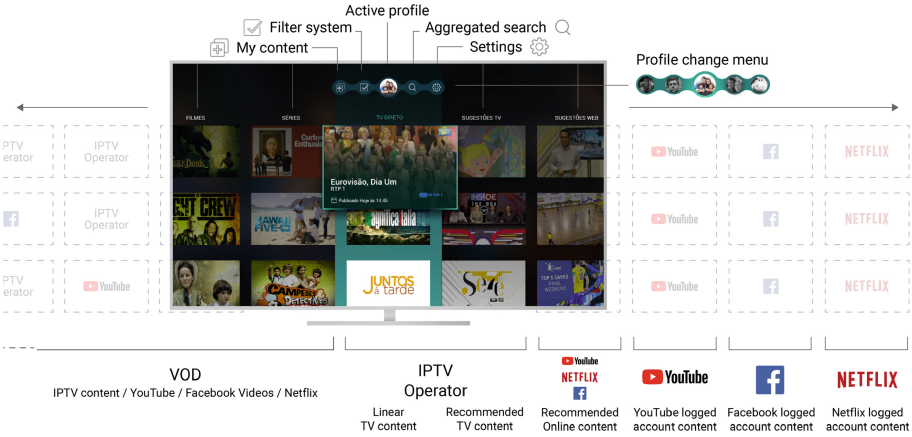

Fig. 2. UltraTV home screen interface architecture

For the assessment of the final version, as the culmination of all the previous developments, the goal of the team was to build a solution that allowed for the unification of TV content, aiming to enhance the entertainment experience offered to different consumer profiles, including younger generations. On the one hand by providing content beyond the traditional broadcast channels to the classic TV consumers

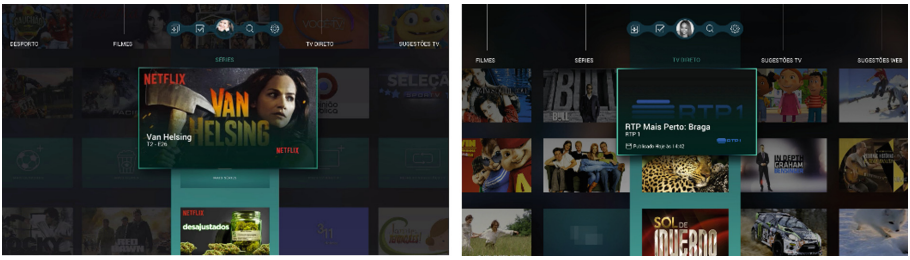

Fig. 3. UltraTV home screen interface with the active focus on the first card of the navigation grid (left).and with active focus on pea menu on top to change profile (right).

Fig. 4. UltraTV full screen interface with the active focus on the contextual menu (left) and with active focus on the cross-reference content on the side menu (right).

(promoting the diversification of their choices). And on the other hand, by aiming to regain 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 on their computer and mobile devices. By targeting a different approach to zapping the prototype fosters the discovery of content from different players and sources, based on the user's profile and tracked behaviors.

3 FT Evaluation of the UltraTV Prototype

A FT plays a crucial role in validating a product, allowing it to be tested in conditions close to the real context of use, revealing potential problems that would not appear in controlled environments like the tests in laboratory. Therefore, a FT is an essential UX evaluation stage integrated in the iterative process for design and development of a product or service [13]. The tests with users (in laboratory and the field) followed the CUE (Components of User Experience) model [14] as a framework to evaluate the components of UX (based on four key dimensions, namely visual/aesthetic, emotion, stimulation, and identification [15–17]. As proposed by the team's previous work [13], a combination of qualitative data-gathering combined with a triangulation of specific validated UX instruments were applied, namely the SUS [18–20], SAM [21] and AttrakDiff [22] questionnaires. Qualitative data was also collected through interviews to gather complementary opinions [23].

3.1 FT Objectives and Procedures

A field evaluation spanning 20 days was carried out in the domestic context. The UltraTV FT took place in 2018 from 12 to 31 of January, using a total of 20 STBs installed at participants' households (see Fig. 5).

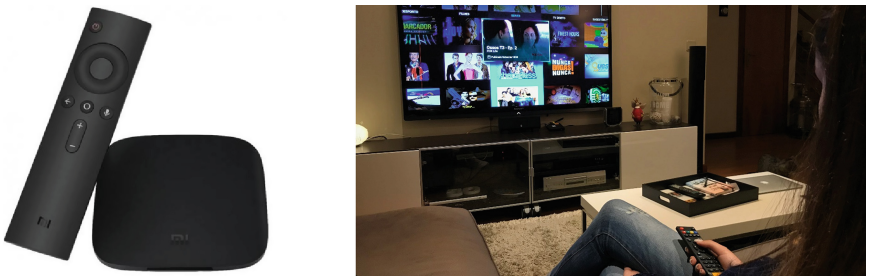

Fig. 5. Xiaomi MI Box 4 K HDR Android TV with Bluetooth remote control (left). A participant using the UltraTV STB in his house (right).

The home-installed test version (see Fig. 5) was an updated version of a previous prototype tested in the laboratory. In this sense, the FT evaluation had the following objectives: (a) Validate the technical solution for the concept of content unification; (b) Evaluate the UX of the prototype; (c) Validate the keys on the remote control used to interact with the system; (d) Evaluate the look & feel of the UI; (e) Validate specific system features (e.g. menus, cross-references, iconography, second level categories in the grid 'silos'); (f) Evaluate the personalization of experience (profiles with personalized grid and suggestions); (g) Evaluate the level of interest and appropriation of the system by the participants.

Some requirements were defined to be a participant in the FT, namely the need to have at least one TV set at home; willingness to use the prototype to access audiovisual content over the test period; have an internet connection at home; and be a regular viewer of online videos from sources like Netflix, YouTube or Facebook. Although only 20 STBs were available to be installed at different households, each STB supported up to 5 users, so participants were asked to invite their family members to join, which allowed the number of participants to rise to 48. This preparation phase (see Fig. 6) ended with the administration of a characterization questionnaire to know more about TV consumption habits.

The implementation phase (see Fig. 6) started with providing each evaluator with a trial package containing one STB, one remote control and an informative brochure. The participants just had to connect the STB to their TVs, set-up the Wi-Fi connection start using the system. The assessment phase also took place during this period (see Fig. 6): Google Analytics was set up to log all interactions on the prototype, by mapping the paths each participant performed in real time. Additionally, two different assessment moments using online questionnaires were applied, being one a halfway survey and the other the final one, including the UX scales SUS, SAM and AttrakDiff. A phone interview, for qualitative data collection, was conducted one week after the end of the FT.

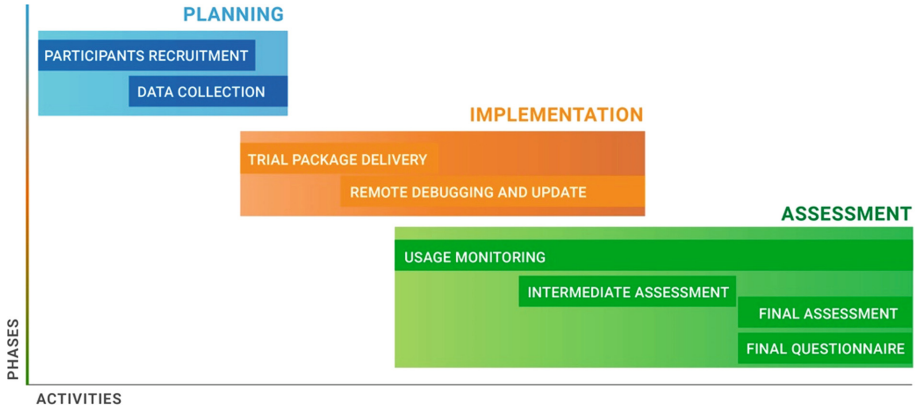

Fig. 6. Phases and activities of the FT dynamization

3.2 Sample Characterization

Prior to the FT, a characterization questionnaire was filled by all participants. The sample included students and power users, ensuring viewers from different age groups, especially target audiences of the UltraTV project – "Millennials" and "Generation Z". From the total of 20 STBs that started the FT, 12 STBs in 12 different houses were selected to be analyzed accounting for a total of 26 individual profiles and 9 family profiles. 8 STBs were not considered due to not complying with using the system for at least 6 h and registered more than 200 interactions. The 26 valid participants included 15 male (57.69%) and 11 female (42.31%) evaluators, aged between 12 and 54. Considering the completed level of education, high school education (38.46%) was most common, followed by a BA (26.92%) and MA (19.23%). According to an age stratification with 5-year intervals (see Fig. 7), the sample was most illustrative of four groups: 26.92% between 20–24 years old and 19.23% between 15–19 years old. These are followed by ages between 30–34, and between 35–39 years. For data analysis purposes participants were also grouped according to their family structure: (a) Families - 10 participants (38.46%); (b) Couples - 8 participants (30.77%); (c) Roommates - 5 participants (19.23%); (d) Alone - 3 participants (11.54%).

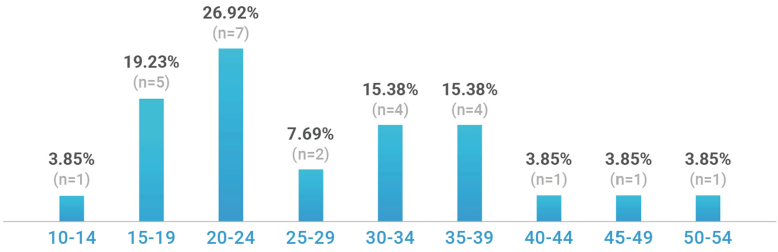

Fig. 7. FT sample by age groups

4 Results and Discussion: Insights from Younger Generations

4.1 Audiovisual Consumption

Based on the Google Analytics collected data from 26 users it was possible to compare the global video consumption stratified by the genres/sources available in the UI grid. Data from the consumption of live TV content and VOD of the last seven days of the associated IPTV service was combined to allow to measure the traditional TV content consumption and compare it to OTT (Netflix, YouTube and Facebook sources) consumption (see Figs. 8 and 9). Considering the viewing time of television content (88.17%), Google Analytics results (Fig. 8) show that the majority of participants spent more time watching linear TV (55.37%). Figure 8 shows that, despite the popularity of OTT sources consumed daily through various devices, as mentioned by the participants in the characterization questionnaire, only 11.83% of the time was spent watching these sources, being most of this consumption dedicated to Netflix videos (10.47%).

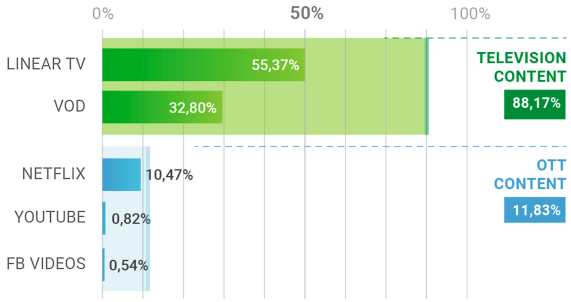

Fig. 8. Time spent watching TV per content source

However, considering that YouTube and Facebook videos are shorter, the number of videos viewed were also analyzed, showing a slight increase in both sources (see Fig. 9). Taking into account that access to Facebook videos and YouTube's subscriptions and recommendations implied the configuration of users' personal accounts, it is considered that this could have been an influential factor for the lower consumption. Besides the fact that the system is connected to a television shared by several members of the household, which was mentioned by some of the participants through the telephone interview conducted after the end of the test period.

Considering the age distribution, the three users with the highest consumption of OTT sources were between 12 and 20 years old. The youngest participant watched exclusively content from Netflix and the participants with 20 years old dedicated an average of 70% of the consumption to internet sources (Facebook and YouTube). By contrast, the participants that only watched regular TV content were between 20 and 34 years old. There were only two cases of participants (19 and 20 years old) with a balanced consumption of TV and OTT content.

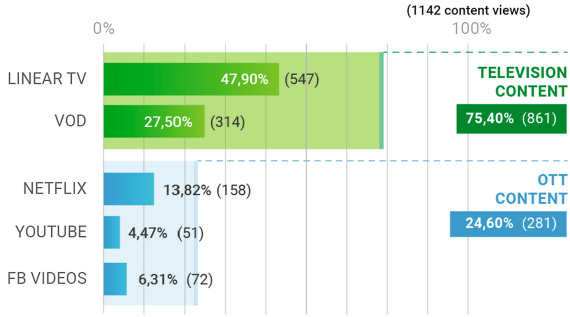

Fig. 9. Number of videos watched per content source

Considering the visualization groups, STBs associated with the "Family" and "Couple" groups achieved the highest total consumption means. Several participants preferred to use a generic family profile rather than their individual profile. These factors can be indicators that the UltraTV application promotes a collective consumption dynamic, a hypothesis acknowledged during the semi-structured interview conducted after the end of the test period. Another factor mentioned in the interviews that influenced the dismiss of individual profiles was the perceived lack of efficiency of the recommendation system (since it needed more time than the test period to make the content suggestions more accurate and more noticeable on the grid).

4.2 User Experience Metrics

The three quantitative UX scales adopted to evaluate the UltraTV prototype in the field were the SUS, SAM and AttrakDiff. Figure 10 provides an overview and highlights of the individual lowest and highest scores (SUS and SAM) according to age group:

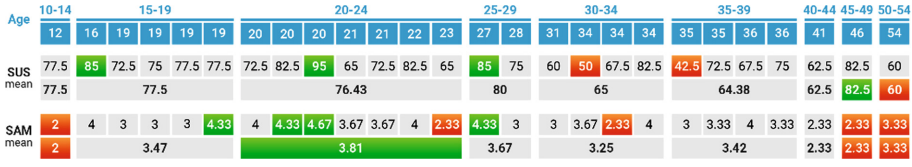

Fig. 10. SUS (0 to 100) and SAM (1 to 5) scores according to age groups

The System Usability Scale questionnaire (based on ten attitude items) addresses the global usability of the system using a Likert scale, regarding controllability, effectiveness, learnability. The global SUS score of 72.4 (from 0 to 100) points out the product as "Good" (Bangor et al., 2009). In addition to the overall SUS score, it was considered relevant to analyze the polarization of opinions for the different items (scores from 1 to 5) according to the participants' individual scores, to verify if their evaluation was related to age or viewing group patterns, associated with specific audiovisual consumption habits.

The positive item #4 (“I think that I would need the support of a technical person to be able to use this system”), #8 (“I found the system very cumbersome to use”) and #10 (“I needed to learn a lot of things before I could get going with this system”) obtained the more favorable opinions. Thus, suggesting that there was a strong conviction in considering the system easy to learn and use.

The only negative scores were pointed by two participants, 35 and 36 years old, in the item #2 (“I found the system unnecessarily complex”) and #6 (“I thought there was too much inconsistency in this system”).

In an individualized perspective, the highest SUS score of 90 points was given by a 20-year-old evaluator, integrated into the group “Roommates”, followed by two evaluators aged 27 and 16 respectively, from the “Family” group, which scored 85 points (see Fig. 10). On the other hand, the worst scores (42.5 and 50 points) were attributed by two young adults, with 35 and 34 years old respectively (see Fig. 10). The first integrated into the group “Couples” and the second in the group “Alone”.

In terms of hedonic and emotional aspects, the Self-Assessment Manikin, as a non-verbal pictorial questionnaire, measured the “satisfaction”, “motivation”, and “sense of control” associated with a person’s affective reaction. Among the three parameters (from 1 to 5), the “sense of control” got the highest score (3.77), followed by “satisfaction” (3.23) and “motivation (3.12).

Regarding age groups, the lowest SAM score (1 point) on the “satisfaction” and the “motivation” parameters was given by the youngest evaluator of the sample (12-year-old) within the “Family” group. Another participant with 19-year-old, from the “Alone” group, also attributed 1 point to “motivation” and by contrast, gave 5 points to the “sense of control”. When comparing with the qualitative data from the phone interview, this participant claimed to have no control over the recommendations, so it is possible that this score results from a confusion in the use of the SAM scale, which inverts the order of the figures in the parameter of the “sense of control”. In this case, it means that this evaluator would have attributed the worst score (1 point) in two parameters related to emotional impact (having also assigned a low score of 2 points in the “satisfaction” parameter), reinforcing how the frustration of failing expectations can have a huge negative impact on the perception of a product. In contrast, the “Roommates” group got the best average (4.07), which may again suggest that the UltraTV system is promoting a collective viewing dynamic regarding the TV screen.

Looking at the SAM scores, the age group between 20 and 24 years was the one with the best means (3.81), followed by the age range of 25 to 29 (3.67). The highest individual mean score (4.67) was rated by a 20-year-old participant within a “Roommates” group. The second-best scores (4.33) were given by two evaluators with 19 and 20 years old, from the groups “Family” and “Roommates”, all of them with an in-depth knowledge of OTT platforms and regular consumption of video content on multiple devices, specially the laptop. Therefore, there is also a willingness to use the system on behalf of these type of users because it provides access to OTT content, with personalized recommendations that can be watched on the television in the shared environment of the living room. As clarified in the phone interviews, this entertainment opportunity related to the TV is therefore clearly apart from the time spent working on the laptop, being perceived as a break time to leisure and relaxation.

As an extension of the SAM scale, the perception of the relational, emotional and aesthetic impact that a product causes on the user is becoming more relevant in the development of systems. For this reason, the AttrakDiff questionnaire, based on pairs of opposite adjectives that measure the pragmatic, hedonic (at the level of stimulation and identification) and aesthetic dimensions, constitutes a fundamental complementary tool to other UX metrics. Due to the limitation of the AttrakDiff platform[3] to 20 evaluators a downsampling process from the 26 participants was required. For this, the inclusion criterion was to be a regular user of the STB (at least 6 h and more than 200 clicks, in addition to filling all the data collection instruments). Aiming to a comparative analysis based on the age, two groups were defined: 10 participants with less than 25 years old and the remaining 10 over that age (see Fig. 11).

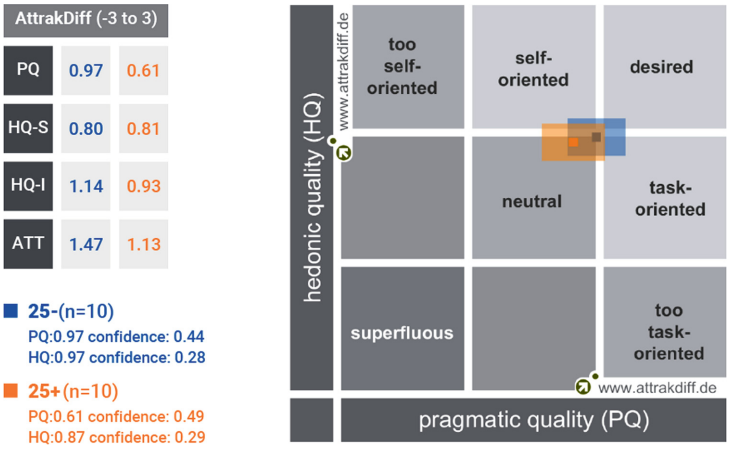

Fig. 11. AttrakDiff comparison of age groups: global scores (left) and confidence rectangles (right).

Similar scores between the two groups were achieved in the "stimulation" dimension (HQ-S), whereas the "aesthetic" dimension (ATT) was the best quality scored by both groups, although significantly higher in the under 25 group. The participants over 25 years old pointed the lowest scores to the "pragmatic" dimension (PQ 0.61) while the younger evaluators gave the lowest ratings to the "stimulation" dimension (HQ-S of 0.80). Considering the distribution by pragmatic and hedonic qualities (see Fig. 11) the UltraTV system was better received by the younger participants. The group under 24 pointed the product towards the "desired" quadrant, while the older group presented a more neutral and contained stance in terms of scores. Both groups display a similar confidence interval. Although younger participants have a best impression regarding the pragmatic qualities, since the system provides suitable answers to tasks integrated in their daily routines (e.g. OTT audiovisual consumption)

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

compared to the other group that did not consider the system as functional regarding their consumption habits ('scheduled viewing' in contrast to 'content discovery'), furthered clarified by the phone interviews.

4.3 Motivational Factors Towards a Content Unification System

The qualitative feedback provided by the participants was also valuable to understand the reasons for their scores and opinions.

Unification and Profile-Based Personalization Towards Content Discovery
The results obtained by all the instruments applied throughout the FT allowed drawing a profile of each user's consumption and identifying the most critical aspects that were addressed in the phone interview. In addition to trying to clarify some of the answers obtained in the other instruments, the participants had the opportunity to express in the interview some additional opinions, namely about unification, profile-based personalization and content discovery.

As for unification, in overall, this was a valued feature because it allowed the aggregation of content, its display and access using a simple UI and cohesive UX [*#118 "The design is simple, good and intuitive. It's easy, and it groups things that are scattered across other platforms"* (age 20)/*#131 "The system alerted me to things I had no clue – for instance, the part of Netflix. I had never used Netflix. Being able to access via this system to Netflix contents opened a door there that I had not explored. And then things like Facebook and stuff … things I would never see. If it were on my phone, I would miss it. And by being there on the home screen, I create that curiosity, and it is easier for you to see the contents"* (age 33)]. However, the consumption of OTT was lower than it was expected [#116 *"Because in class we are always online and on Facebook when we get home, we want to stop a little. I've been watching more movies and live television than Facebook and YouTube"* (age 19)], which can be explained with some limitations in the prototype regarding Facebook and YouTube pairing, that may have dissuaded some participants, and also some problems in the quality of the video streaming loaded from the IPTV provider application (paralyzed, blurred, pixelated image), which in some cases led to less regular use of the UltraTV STB [#173 *"I was never able to log in on Facebook. And on YouTube, the streaming was very bad"* (age 35)].

Also, the fact that not all users have paired their Facebook and YouTube accounts to their profile has also conditioned their UX for accessing OTT content [#128 *"I am not the type of user who feels the need to use social media on the TV set. That is, I do not see this profile need. And usually, here at home, we will all watch the same. I do not feel this need to change profiles depending on who is watching television. I would say this functionality will be most useful when our children grow up"* (age 37)/#117 *"I didn't notice which profile I was using. Usually, we were all together, or someone was already watching. As for me, I used the system together with everybody, and, it worked out well as it ended up improving a lot of suggestions that helped to find stuff, we all liked"* (age 22)]. Among participants who connected their accounts, some said they lacked native features of other platforms, such as commenting, liking, or sharing. Other users also mentioned privacy issues, highlighting the need for a PIN to access the

profiles, which would also safeguard the personal social media accounts. Others indicated that the "family" profile fostered the discovery of content and collective viewing, in some cases with a preference for using this general profile rather than the individual profile *[#126 "I think it is because of our age group… Because we are students and we live together, we end up having all the same interests. So, we usually watch television together, or else the programs we see fit all within the same genre. But in our case, as a group of friends, we ended up opting to use a more general profile that ends up being shaped to our tastes"* (age 21)/*#126 "Just because this box has profiles and our operator doesn't have it, is already an advantage. And in your case, it helped us, for it has aggregated the content. And in the case of our operator's box, we have to actively search the contents to have references of things that we want to see, instead of having already some suggestions available"* (age 21)].

The validation of a content discovery proposal, fueled by the aggregation of TV and OTT content at the same level and using a profile-based recommendation system, suffered both from the regularity of the usage and the short duration of the test (20 days). If, on the one hand, the ease of access and automation of suggestions was recognized as an interesting premise *[#106 "This is great for me. I do not want to walk through looking at two hundred channels instead of what I am really interested, which is like… 10%.*" (age 36)/#114 *"And also by showing related videos will also display shows that, for example, you would not discover if it were not so"* (age 36)], on the other hand, the lack of more accurate suggestions affected the expectations of the users [*#173 "I currently think the prototype has a skewed limitation of my preferences. But I think that's what a more advanced version would have improved and would encourage me to want the product. If you had the discovery of content with the search feature that would be great … and perhaps fine-tune the related content on the right-side menu"* (age 35)]. Some users missed features to directly assign what you like to watch to help the recommendation system. This demonstrated that the context menu option of "I do not want to watch this content" went unnoticed to some participants, either because they did not find it or because they did not understand that by activating it would affect the recommendations on their profile-based grid.

Audiovisual Consumption in Other Devices

Although the UltraTV consisted of a television ecosystem, based on several consumer-oriented devices (television, mobile devices – smartphone and tablet – and computer), the prototypes tested focused exclusively on the STB version of the system. This option considered the TV consumption at home the most relevant use case scenario for the unification of content and UX proof of concept. Namely, regarding the acceptance of a disruptive approach to television consumption, and potential penetration in its context of excellence: the domestic leisure environment of the living room (simulated in previous laboratory tests and field tested in participants' homes). At the same time, the reconciliation with the television screen, to the detriment of the daily and ubiquitous consumption in the other devices, was considered another relevant issue to be validated. In particular, younger audiences tend to be further away from television in favor of the use of personal equipment such as the computer and mobile devices. Therefore, we tried to find in the initial, intermediate and final questionnaire and, in the last phase by phone interview (see Fig. 6), if the use of UltraTV would potentially have an impact

on the consumption of audiovisual content in other devices. In particular, it was intended to verify if the participants anticipated a reduction of the consumption in these devices, privileging the visualization of videos on the television screen.

The participants highlighted the advantage of using the TV for collective viewing in a relaxed environment, compared to other devices such as the computer, more related to work and study tasks. The comfort and quick access to a range of recommended content on the home screen grid using only one-to-two-click (s) after connecting the TV/STB, without having to go to the browser or access to apps, were mentioned as advantages and motivations to reduce the consumption of audiovisual content at home through other devices.

In any case, a cross-platform solution was also considered to be relevant, especially for mobile use. Some participants also suggested the advantage of using mobile devices as a second screen to search for additional information about content without having to interrupt their viewing or even search before deciding to view it. Part of the prejudice of some participants, regarding the use of television, has to do with the limitation in the quantity, diversity and even up-to-date content available. Younger users continue to look at the television offer as old fashioned, assuming that the content programming provided by operators and channels will not match their tastes, compared to the ease of downloads made on the computer. The aggregation of OTT content, together with an integrated search and the personalized recommendation of contents that UltraTV offers, are precisely aimed at countering this view. This type of opinion confirms that there is a need for a product that unifies current and personalized content in a way that appeals to the most skeptical audiences about TV, especially the younger generations.

Willingness to Adopt UltraTV

The final questionnaire included a last closed question to perceive whether the participants were willing to change their current TV service for a commercial and fully functional version of the UltraTV system at the same price point, followed by an open-ended question to justify their choice. 76.92% of the participants replied they were willing to do so, which suggests a good indicator of the willingness for the unification concept and its validation.

Of the 26 respondents, only 6 participants – with 16, 34, 35 and 36 years – stated that they did not want to adopt the system. This result confirms the other UX scores that point to young adults (in their 30s) as the most skeptical audience to change their routines in favor of new products. The younger (16 years old) participant who stated that he did not want to adopt the system justified the answer by claiming that he was satisfied with his current service without further mentioning negative issues about the UltraTV prototype.

Regarding the negative issues, the need for more control over the system and the lack of advanced features and settings were considered the reasons to not adopt the UltraTV application if available on the market. The lack of control refers to some navigational difficulties that can be related to the disruptive interaction model of this system compared to regular services (namely the use of a grid with the use of directional keys without a channel alignment or numeric keys on the remote control). And also non-intuitive remote-keys for the activation of some specific areas (e.g. access to the side menu in full screen). The lack of other features and settings may also have been

mentioned because some sections of the prototype were not fully functional during the test. This was the case of the lack of the "Search" feature, which did not allow quick access to channels and to browse specific content. Still, this qualitative set of data is in line with the quantitative results obtained in the SUS, SAM and AttrakDiff scales.

Finally, the feedback gathered in the phone interviews also allowed to clarify some of the results. The most negative issues referred by less than a third of the evaluators were: (a) the limited number of features as compared with the service they currently have and are satisfied with; (b) the lack of a chronologic channel surfing feature, usually called EPG (Electronic Program Guide); (c) some resistance to the organizational logic of the content on the grid. It should also be noted that users who are frequent consumers of TV, independently of the age, showed higher resistance to the regular use of this system.

5 Conclusions

The emergent iTV ecosystem leans towards on-demand content without dismissing the viewing of the shared linear experience on the TV screen, complemented by flexible access across mobile devices. Recent releases in the iTV industry confirm the growing expectations towards personalization and unification features, favoring the aggregation of sources, merging linear with non-linear videos. Despite some unification solutions released on the turn of 2018, the single UI horizontal unification embodies a different path from the prevailing offers that provide content in proprietary apps.

Acknowledging these challenges, the UltraTV project aims to enhance the entertainment experience by focusing on unification and personalization features. After several iterations assessed by experts and users in laboratory environment, the quantitative and qualitative UX results from the final prototype, tested in the field, reveal significant insights regarding younger generations, often detached from the TV screen and traditional TV content in favor for other devices and formats.

The outcomes from this empirical research sustained the motivation of users to have access to an iTV solution that includes content beyond the traditional broadcast channels, upgrading the offer with OTT content. Positive feedback was higher in young generations, across all instruments further corroborated and detailed by qualitative data from interviews and open-ended questions from a final questionnaire. This data provides interesting perceptions suggesting that unified iTV applications may be significant in retaining younger generations regarding the TV screen, including the cord-cutters. Nevertheless, there were limitations inherent to the trial procedure, namely the fact that not all designed features were fully functional, as well as API restrictions which prevented the inclusion of native features from Netflix, YouTube and Facebook. Despite these limitations, the validation of the unification and personalization concept, along with the recommendations provided by the participants, constitute an encouragement to further development and testing with more advanced functionalities. These may include content discovery and social networking tools embedded in the system, and also available on second screen devices to be used both as a remote control and a way to interact with complementary information.

In addition, the UCD methodological approach adopted by the UltraTV project, that recognizes industry trends and integrates the feedback of end-users in different stages of the UX evaluations using quantitative and qualitative instruments, may provide valuable contributions regarding frameworks for product development as well as user's insights agnostic to players' interests to push the next generation of television systems.

Acknowledgments. This paper is a result of the UltraTV - UltraHD TV Application Ecosystem project (grant agreement no.17738), funded by COMPETE 2020, Portugal 2020 and the European Union through the European Regional Development Fund (FEDER). Authors are grateful to the project partners.

References

1. Ericsson-Consumer lab. TV and Media 2017: A Consumer-Driven Future of Media. Ericsson, Stockholm (2017)
2. Abreu, J., Almeida, P., Teles, B.M.: TV discovery & enjoy: a new approach to help users finding the right TV Program to to Watch. In: Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video -TVX 2014, pp. 63–70. ACM, Newcastle (2014)
3. 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**(1), 57–74 (2017)
4. Vanattenhoven, J., Geerts, D.: Broadcast, video-on-demand, and other ways to watch television content. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX 2015, pp. 73–82. ACM, Brussels (2015)
5. Nielsen-Company. Screen wars: The battle for eye space in a TV-everywhere world. Nielsen Insights, Nielsen, New York (2015)
6. Almeida, P., et al.: Iterative user experience evaluation of a user interface for the unification of TV contents. In: Abásolo, M.J., Abreu, J., Almeida, P., Silva, T. (eds.) jAUTI 2017. CCIS, vol. 813, pp. 44–57. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90170-1_4
7. Almeida, P., et al.: UltraTV: an iTV content unification prototype. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX 2018. ACM, Seoul (2018)
8. Lowdermilk, T.: User-Centered Design: A Developer's Guide to Building User-Friendly Applications, vol. 1. O'Reilly Media, Inc., Sebastopol (2013). (M. Treseler, ed.)
9. Abreu, J., Almeida, P., Silva, T., Velhinho, A., Fernandes, S.: Using experts review to validate an iTV UI for the unification of contents. In: Proceedings of The 12th International Multi-Conference on Society, Cybernetics and Informatics (IMSCI 2018). Springer, Orlando (2018)
10. Almeida, P., Ferraz de Abreu, J., Oliveira, E., Velhinho, A.: Expert evaluation of a user interface for the unification of TV contents. In: Proceedings of the 6th Iberoamerican Conference on Applications and Usability of Interactive TV - jAUTI 2017, pp. 59–70. Universidade de Aveiro, Aveiro (2017)
11. Almeida, P., Jorge, A., Fernandes, S., Oliveira, E.: Content unification in iTV to enhance user experience: The UltraTV Project. In: Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video - TVX 2018. ACM, Seoul (2018)

12. Waterman, D., Sherman, R., Wook Ji, S.: The economics of online television: industry development, aggregation, and "TV Everywhere". Telecommun. Policy **37**(9), 725–736 (2013)
13. Abreu, J., Almeida, P., Silva, T.: A UX evaluation approach for second-screen applications. In: Abásolo, M.J., Perales, F.J., Bibiloni, A. (eds.) jAUTI/CTVDI -2015. CCIS, vol. 605, pp. 105–120. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-38907-3_9
14. Mahlke, S., Thüring, M.: Studying antecedents of emotional experiences in interactive contexts. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 2007, pp. 915–918. ACM, San Jose (2007)
15. 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
16. Hassenzahl, M.: User experience (UX): towards an experiential perspective on product quality. In: Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine on - IHM 2008, pp. 11–15 (2008)
17. Hassenzahl, M., Tractinsky, N.: User experience - a research agenda. Behav. Inf. Technol. **25**(2), 91–97 (2006)
18. Bangor, A., Kortum, P., Miller, J.: Determining what individual SUS scores mean: adding an adjective rating scale. J. Usability Stud. **4**(3), 114–123 (2009)
19. Brooke, J.: SUS - a quick and dirty usability scale. Usability Eval. Ind. **189**(194), 4–7 (1996)
20. Martins, A.I., Rosa, A.F., Queirós, A., Silva, A., Rocha, N.P.: European Portuguese validation of the System Usability Scale (SUS). Procedia Comput. Sci. **67**, 293–300 (2015)
21. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment semantic differential manikin and the semantic differential. J. Behav. Therapy Exp. Psychiatry **25**, 49–59 (1994)
22. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität, pp. 187–196. Vieweg+Teubner Verlag (2003)
23. Patton, M.Q.: Qualitative Research and Evaluation Methods Qualitative Inquiry, 3rd edn. Sage Publications, Thousand Oaks (2002)