Taylor & Francis
Taylor & Francis Group

User interface guidelines for the control of interactive television systems via smart phone applications

Regina Bernhaupt[a]* and Michael M. Pirker[b]

[a]*ICS, IRIT, Toulouse, France, IRIT, Group ICS, 118 Route de Narbonne, 31062 Toulouse, France;* [b]*Ruwido Austria, Koestendorfer Str. 8, 5202 Neumarkt a. w. Austria*

There are a growing number of smart phone applications allowing the user to control their television, set-top box or other entertainment devices. The success of these applications is limited. Based on findings from media studies in Austria and France focusing on how people currently use their TV and iTV systems and associated devices, this article describes recommendations for the design of a smart phone application enabling users to control Internet Protocol Television (IPTV) systems including all connected entertainment devices. Recommendations include the need to allow users to control devices that are related to the IPTV experience (not only the set-top box or television set) and the focus on scenarios of usage like supporting listening to music, enjoying a movie or controlling the connected home. Based on similarities and differences found in the two samples, future smart phone applications for controlling TV will only succeed if they provide meaningful functionalities that satisfy the (varying) user needs, support personalisation and personal usage and respect the limitations of mobile phones with respect to possible parallel activities performed.

Keywords: IPTV; mobile telephone; usability; user interface guideline; control; smart phone application

Introduction

Watching TV is still enjoying great popularity as a leisure time activity. While there have been claims that the Internet, personal computers (PCs) and mobile devices will replace traditional TV, the current trend is rather to (still) watch TV while using other devices at the same time.

What we found in a series of studies, focusing on the domains of TV, interactive TV and Internet Protocol TV (IPTV), is that although people use an additional device during 60% of the TV watching time, the majority (60%) of the TV watching is still a social activity (Bernhaupt *et al.* 2008, 2011). Given this high percentage of parallel usage, secondary devices – and especially mobile devices – seem to be a good candidate to enhance the current TV and entertainment-related activities in the living room.

One of the central problems for enjoying TV in the living room is that the activity of simply watching TV, selecting some kind of media content or setting up any form of entertainment devices has become more complicated than ever before due to the ever-increasing complexity of today's TV and entertainment systems. About half a decade ago, the following steps were sufficient to watch a TV programme on, for example, channel three:

- Step 1: turn on TV (either by pressing a button on TV or by pressing a button on the remote control).

- Step 2: change to channel three (by pressing a button on TV directly or selecting the number button three on the remote control).

In contrast to this, watching TV today in a French IPTV household requires the following steps:

- *Step 1a*: take first remote control to turn on TV.
- *Step 1b*: take second remote control to turn on the set-top box.
- *Step 2*: take TV remote control and change input source to the correct HDMI source by navigating the user interface with the navigation keys (to receive set-top box input).
- *Step 3*: take the remote control of the set-top box and confirm that you want to watch (live) TV by using the navigation keys and ok in the user interface.
- *Step 4*: type in 003 on the numeric keys to select the channel three [or alternatively use the electronic programme guide (EPG) to access channel by using the guide button, navigation buttons and ok].

There is clearly a need for new forms of control in the entertainment environment. One possibility to enhance the current control of entertainment-related devices in the living room is to use smart phone applications to enable the user

---

*Corresponding author. Emails: regina.bernhaupt@irit.fr, Michael.Pirker@ruwido.com

to control a variety of devices and to perform additional tasks that enhance the overall entertainment experience. Smart phones additionally enable the user to use different forms of interaction techniques to control their entertainment environment than with button-based remote controls, like speech, gesture or touch input. Instead of navigating through a set of menus with a set of different (infrared based) remote controls, users might simply talk to the mobile phone, perform a gesture, or interact with the smart phone application via touch input.

The goal of this article is to provide a set of design recommendations for the next generation of smart phone applications that enable users to control all interactive TV/IPTV systems and connected devices in their homes with one single device in an efficient and effective way, while at the same time providing an improved user experience compared to a standard remote control based button input. The focus of the recommendations is *only* on *how to enhance the direct control of entertainment devices*, not on other ways of enhancing the TV watching experience (like providing enhanced versions of electronic programme guides, second screen applications accompanying the programme like quiz shows, etc.). The set of recommendations is based on media studies in France and Austria using the ethnography-based playful probing approach. The studies investigate how people currently use their smart phone as part of their entertainment environment, what ideas users have for future usage of smart phone applications with the research focus to understand what type of problems related to interactive TV/IPTV can be successfully solved with smart phone applications, and finally how to enhance the overall user experience of today's IPTV offers with such a type of application.

State of the art

TV and interactive TV in France and Austria

Watching TV is a central activity in French and Austrian households. In France the daily TV watching time for people 15 years and older in the year 2012 was four hours and six minutes (France Mediametrie 2012) and in Austria it was two hours and 49 minutes for people 12 years or older in 2012 (ORF Media Research Austria 2012a). The distribution of TV reception differs in the two countries with France offering the highest distribution of IPTV offers for households (approximately 25 million individuals in 2012, which is about 42% of the TV population of 59 million individuals in France Metropolitaine – cf. France Mediametrie 2012), while Austria is having only one major IPTV provider with less than 1% market share. The majority of households in Austria receive their TV signal via satellite, cable TV or DVB-T (terrestrial digital video broadcast). In the year 2012, 55% of Austrian households receive their TV signal via digital satellite reception, 40% via cable TV and 6% use digital terrestrial TV (DVB-T) only (ORF Media Research Austria 2012b).

Overall, the two countries have different market situations, where the diffusion of IP-based television services is fairly different. Whereas in France the diffusion of IP-based television services and Triple Play options for customers is remarkably high with several providers offering these services, the IPTV landscape in Austria is by far not that developed, with only one major IPTV provider. The high percentage of satellite viewers in Austria is caused mainly by the spill-over from German satellite television combined with a strongly regulated TV market in Austria. Additionally, also the topological situation in Austria does not favour the diffusion of IPTV, as broadband Internet outside of larger cities was not easily available and most people already settled to satellite or cable TV before IPTV got available.

(IP)TV and media usage

Gauntlett and Hill (1999) have studied TV usage in the home in one of the largest media studies, including 500 participants. Their key findings were that TV is still a focus of attention and watching TV is influenced by social interaction. Viewing habits differ immensely between age groups, gender and personal interests. Bernhaupt *et al.* (2011) have investigated media consumption in a series of ethnographic studies, reporting that

(1) the living room is still a social and an individual place, and that technology is only becoming part of that personal space if it supports the feelings of being at home, respecting people's privacy,
(2) TV is strongly related to other activities in the home,
(3) interaction techniques, especially the growing number of remote controls, are perceived as difficult to use, and
(4) watching TV is still a social activity supporting people's desire 'to be together'.

Watching TV is principally seen as a leisure activity, and most often it is a rather passive activity (Gawlinski 2003). Several studies have shown that the interactivity of interactive TV systems is not always welcomed, but a series of studies demonstrated that additional services and features can enhance the overall user experience. Ethnographic studies and ethnographically oriented field studies investigated the application of new services and systems in the home, including the application of new recommender systems (Bernhaupt *et al.* 2011), personalisation approaches and the enhancement of the social experience (Cesar *et al.* 2008). Other studies on media consumption typically focused on a particular media technology, for example the introduction of a set-top box in people's home (O'Brien *et al.* 1999). Seager *et al.* (2007) investigated the use of triple-play services, focusing on the different devices employed in the households.

To summarise, media consumption is steadily changing. While five to seven years ago, TV was the main source for

entertainment-related content like movies, series or news, today Internet access and a multitude of new (specialised) devices have changed who is consuming what kind of media on which device. Traditional TV watching is becoming more and more a multi-device (multi-user) activity. Especially in the area of interactive TV, these changes in media consumption are important, as the future generation of iTV systems shall support new forms of media consumption without making the usage of interactive TV system more and more complex.

To support the development of future interactive TV systems and services and their control, there is a need to understand user practices and behaviours.

Controlling the interactive TV system: remote controls

When interacting with interactive TV, the standard way is the use of a remote control. There are several problems occurring when the TV viewers use remote controls. First of all, the usage of a remote control is often limited to a small set of buttons that are frequently used. According to a study by Logan *et al.* (1995) on the television habits and frequencies of use of different buttons on the remote, the buttons 'P+/P-' to change channels made up more than 50% of the use of the buttons of a conventional remote control, the buttons 'V+/V-' to change volume up to 19% and the use of digital buttons is 17%.

Another usability problem is the number of remote controls and the function of each particular remote control. In the living room, nearly each technical device that can be operated remotely has its own remote control. A study from Darnell (2008) on new TV system shows that 'the most serious problems involved determining which remote control to use for a given function, making remote control mode errors, and having difficulty going back to TV from recordings and locked channels'. New TV systems are difficult to use, especially for less-technically-inclined consumers (Darnell 2008).

In Ofcom's (2007) report 'easy to use digital television receivers: remote control buttons and functions used by different types of consumer', Freeman and Lessiter discovered in usability test that participants have a 'preference for controlling television receiver equipment with one remote control', and that almost all participants have buttons on their remote controls whose function they do not know and which they avoid pressing, so-called 'easy to ignore buttons on remote controls that are not used or needed'. They also highlight the importance of functional groupings on remote controls. They identified four functional groupings, which are 'Core operations' (power on/off, volume control, mute, channel entry and up/down), 'Interactive services' (EPG, digital text services, red button interactive services, navigation buttons), 'Recording' (play, pause, record, forward, etc.) and 'Accessibility' (subtitles, audio descriptions).

While on a general basis the problem of having only one remote control can be solved by using a universal remote control (URC), in our experience as remote control producers, the complexity of programming and the number of buttons on such URCs often hinder their successful long-term usage in households by all household members. Other solutions to enable the control of all devices with only one remote control or device are changes in the technological infrastructure, for example by providing a URC framework to be integrated in the software architecture of all IPTV devices (Epelde *et al.* 2009).

Interaction techniques related to mobile phones

Smart phones do support the following three basic ways to enhance the interaction: (1) touch-based interaction, (2) gesture and (3) voice/speech. Other possible ways of interaction (e.g. using the camera) were not the focus of this study, as the study was focusing on 'mainstream' interaction modes that are available on today's smart phones.

Recent research investigated, for example, using hands as the input device (Dezfuli *et al.* 2012), using digital pen and paper approaches and guidelines therefore (Hess *et al.* 2011) or projected interfaces and finger input (Harrison *et al.* 2011). Although promising for future applications, all of these approaches have in common that they are prototypical [e.g. needing markers on the finger and hand of the user and a camera for tracking and position detection in the case of Dezfuli *et al.* (2012) or working only in a PC-based environment with applications translating the input signal (Hess *et al.* 2011)]. Additionally, as described by Hess *et al.* (2011), while offering advantages like customisable interfaces, these prototypes also have disadvantages compared to standard remote controls, e.g. the lack of haptic feedback or the need for bi-manual operation, while also sharing some disadvantages of standard remote controls, like bad visibility in low light conditions. Another issue why these new interaction modalities were not in the focus of this study was simply that they are neither (yet) ready for nor available on the mass market, as opposed to aforementioned interaction modes that are supported by today's smart phones.

Voice/speech

The study of using voice as a means of interacting with a computing system has its root in the area of natural language processing (Furnas *et al.* 1987). Elder (1970) was one of the first researchers to consider speech as an interaction technique. When speech is used for interaction with the interactive TV user, the user can interact hands-free in a natural way (without having any device in the hand). For the interaction with an iTV system two possibilities exist: the user can either use natural speech, or the interaction is limited to a set of commands. Goto *et al.* (2003) have investigated natural speech by conducting a Wizard-of-Oz experiment to investigate natural language for interacting with TV. Based on their goal to design a 'television operation interface easy enough for anybody to use', results were

promising. Participants in the study stated that programme selection was easy and that they could directly access any information necessary without the need to go through hierarchical structured menus or programme information.

Turunen *et al.* (2009) presented a multimodal media centre interface, combining voice, gesture and physical touch. Users were given full control over the media-centre via its speech user interface. Its grammar was containing about 110 words, which reached a high recognition accuracy of more than 93%. In their user experience evaluation, they found that users preferred physical touch and gesture interfaces. Speech input was less preferred by the users – it was considered not robust by the users, which was probably due to the 'small' number of words. Turunen *et al.* explained the low preference rating with the fact that the multimodal interaction with gesture and vocal might have disturbed the results of the vocal interaction (e.g. the microphone was in a 'more challenging' angle).

The second approach for speech is the usage of voice as a triggering sound. In their study, Igarashi and Hughes (2001) present three different types of what they call 'Voice as Sound' (VaS) technique. The first one is to control functions induced by continuous voice, meaning the user has to keep making a sound (e.g. '*Volume up, ahhhh*') for the command to continue; in this case, the volume of a TV set continues to increase while the '*ahhhh*' continues. The second type is an extension of the first one, where the pitch of the voice is used as a one-dimensional lever or slider. For instance, if the user is using '*Volume up, ahhhh*' and is raising the pitch of his '*ahhhh*', the volume will increase faster. Igarashi and colleagues also combined this technique with a speed-dependent automatic zooming interface (Igarashi and Hinckley 2000). The last type of VaS technique is a slightly different one, changing the input modality from continuous parameter control to discrete value selection, which Igarashi and Hughes (2001) called 'tonguing'. The objective was to offer an alternative for non-continuous variables, such as channel, for instance. This technique simply detects discrete peaks in sound; an example for voice stated by the authors would be '*channel up, ta ta ta*' to increase the channel number by three. Alternatively, these peaks in sound can be triggered by users simply clapping their hands or snapping their fingers instead of using short voice-noises. The advantage of these techniques is that the users have immediate, continuous control and can stop commands as soon as they are satisfied, because of the immediate feedback. Further advantages of these techniques stated by the authors include language-independency and simplicity.

While such a type of interface can be helpful for users that cannot use their hands due to physical disabilities, the participants in the evaluation study were less positive. This type of interaction has been perceived as unnatural.

Voice/speech interaction per se is in general well perceived by users, but it poses a set of difficulties when applied in the context of iTV with the main goal of enjoying media. Integrating sound/speech recognition in this environment is difficult; it implies the necessity that the user actively activates the speech recognition to avoid any misleading interpretation of the ambient noise or sound in the living room. Apart from the fact that speech recognition accuracy and recognition of certain languages/dialects is still an issue, it additionally needs enhanced technical solutions (e.g. special microphones in TV or remote control) to enable the user to perform speech input. On the other hand recent findings suggest possible enhancements by providing users, for example, naturally phrased fill-in-the-blank speech examples that help guide the user (Stifelman *et al.* 2013).

Gesture

Gesture has become widely accepted as a possible means of controlling devices due to its successful use in video games. It offers the advantage that it enables the user to solely focus on the TV screen, as it does not require visual attention from the user on the remote control, but supports a completely blind usage. Gesture interaction has been reported to be well accepted by users (Kela *et al.* 2006, Pelling *et al.* 2009). What is important to note is the question whether or not mobile phones can be used as pointing devices reasonably well. Research evidence here is contradicting, arguing that pointing with mobile phones is not possible (Lin *et al.* 2010), while others have shown the ability to use them for pointing (Pelling *et al.* 2009).

Gesture interaction based on a Wii-mote has been recently investigated in the context of TV (Bailly *et al.* 2011b) replacing traditional remote controls. Findings indicate that users naturally perform pitch and yaw gestures when using gesture interaction instead of using a remote control. Gestures seem appropriate to enable eyes-free interaction and to avoid the continuous problem of selecting buttons on a standard remote control. In a second study, the same authors studied freehand gesture. It was shown that the optimal mapping between midair gestures and directional actions strongly depends on context (Bailly *et al.* 2011a). Using linear, marking and finger-count menus, the authors achieved an accuracy rate above 93% in their user tests. In their conclusion they state that typing on a physical remote control will remain faster than using gesture to control the iTV with freehand gesture. However, users mentioned that they would appreciate to be able to use freehand gesture as a 'complementary tool' for specific actions.

Hybrid techniques combining midair gestures and buttons were especially efficient, with the further advantage of compatibility with pure button-based techniques. As designers will want to look for eyes-free interactions, techniques using a directional pad or pitch and yaw for transposing 2D marking menus to the midair space are recommended. However, purely gestural techniques tend to be slightly lower in terms of performance. It is necessary to focus on keeping it simple, and to keep in mind that designing for the entertainment in the home, especially for the living room, poses a unique context.

Touch

Robertson *et al.* (1996) as well as Enns and Mackenzie (1998) have already suggested enhancing or replacing remote controls with touch-based input. Robertson *et al.* investigated the usage of a PDA as a companion device to operate iTV services and developed guidelines for multi-device applications, while Enns and Mackenzie presented a remote control including touch and gesture input and described advantages of touch-based remote controls. In general, touch interaction has been becoming more and more important for the mass market with the rise of smart phones. Nowadays, almost everyone is able to interact with a touch-screen almost naturally. While touch interaction was well researched within the last 20 years (Goldberg and Richardson 1993), it still seems to have difficulties to enter in today's living rooms (Pirker *et al.* 2010). The drawbacks of classic remote controls are said to be their lack of scalability and their difficulty in supporting new technologies. Touch-based interaction for remote controls could help to simplify the interaction, supporting flexibility in terms of task-oriented support by simply changing the number of buttons shown on the screen.

Yet mobile phones including touch may not only be used to directly control iTV devices. Tsekleves *et al.* (2007) investigated the usage of mobile phones as second screen devices, where users preferred the PDA prototype over the standard remote control.

Touch interaction can pose a set of usability problems, including the limitation that touch interfaces cannot be used blindly (compared to a standard remote control were users can feel the buttons), that touch interaction needs new forms of communication (protocols) with the TV and iTV system (most standard devices, e.g. TVs, still operate with infrared as input) like connectivity via Wi-Fi (WLAN) or Bluetooth. On the other hand, touch has been proven to provide a better overall user experience when interacting with an iTV system (Pirker *et al.* 2010).

Mobile phone applications and devices for ITV systems

Using mobile phones to control an ambient media player has been described by Lorenz and Jentsch (2010), proposing a set of gestures to control the ambient media player. A complete survey on smart phones as input devices for ubiquitous environments is available in Ballagas *et al.* (2006).

From a user's perspective, there is a variety of smart phone applications currently available for controlling iTV and IPTV systems. All producers of set-top boxes and the majority of I(P)TV providers in Europe, North and South America and Asia offer this type of applications. Also providers of over the top (OTT) services offer such applications, for example, Apple's 'Remote' app,[1] developed for iPhone, iPod and iPad. To a limited extent, such applications are also available for mobile phones with the Android operating system. The limitation of all these applications is

that they typically can only control the set-top box. Once the user wants to control any other device in the living room, it is necessary to use an additional remote control. Smart phone thus only becomes another remote control, but has the advantage that it can be operated independent of its position, as most remote controls are still based on infrared, so users have to point to the device they want to control (line of sight). On the contrary, smart phone applications operate via other technical infrastructure like Wi-Fi, without the line of sight restrictions.

To overcome the limitations of Wi-Fi-based control of only the set-top box, there are a set of devices allowing to control any type of IR-based device in the living room. Examples are L5 or myTVRemote for iPhones.[2] These products are simple add-ons that are attached to the phone and that are able to emit infrared signals. A set of single standing solutions like Peel or Gear4 (Unity Remote) offer the user the possibility to put the device in their home, without the need to plug it on smart phone. Other products additionally allow controlling the whole home infrastructure, including heating or lights, for example BeoLink from Bang & Olufsen.

Devices enabling the control of all different types of technical protocols used for entertainment devices are (to the best of our knowledge) currently not available on the market, neither have they been investigated from a scientific media study perspective.

Current limitations and challenges

The control of interactive TV and its applications is traditionally done with devices that are designed for this purpose, in the majority of cases a standard remote control. These devices have limitations not only in terms of usability (several remote controls are not satisfying when used, not efficient or effective) but also in terms of adaptability (remote controls cannot be easily changed) and scalability (it is impossible to add buttons for every service or feature available on interactive TV). There have been a variety of experimental studies and prototypes offering new solutions for the control of TV: ranging from freehand gesture interaction to usage of the palm as input device. None of these solutions is currently ready to be deployed at the mass market. The focus of our study thus was to look at currently available technology in the households (smart phone) and investigate how the three most common interaction techniques related to mobile phones (touch interaction, gesture and voice/speech) can be used for the control of interactive TV applications.

Study motivation and objectives

This study was motivated by findings within an industrial project called 'Living Room 2020' financed by ruwido austria. The goal of the project was to investigate current media consumption patterns in the living room and beyond to inform the next generation of IPTV user interfaces and

interaction concepts. What we learned in this project is that a number of IPTV providers in Europe and Canada complained that their hotlines have a steadily increasing number of calls, related to the connection of devices and access to various sources of media content, which they are not able to solve any more. One of the currently proposed solutions is to use smart phone applications to allow users to perform their TV- and entertainment-oriented tasks in an easier and more enjoyable way. The success of these applications is limited, most likely due to the fact that they only allow controlling one device at a time, e.g. the set-top box.

The goal of the study was to understand how people currently use media systems and to identify problems and user experiences in IPTV households. Based on these findings, a set of design recommendations to inform the development of a next generation of smart phone applications are developed. The study was guided by the following research questions:

(1) How people use their smart phone as part of their entertainment environment?
(2) What type of problems related to IPTV can be successfully solved with smart phone applications?
(3) How to enhance the overall user experience of today's IPTV offers based on such applications?

Method

Smart phones have become available in almost all households in Europe. While they are used for a variety of applications, it is still not well understood what requirements people have when using smart phones for the control of entertainment-oriented devices or more specifically interactive TV and IPTV systems. We chose an ethnographic oriented approach called playful probing (Bernhaupt *et al.* 2007) − a variation of the cultural probing approach (Gaver *et al.* 2004). The methodological approach has been proven to be helpful in understanding users' habits, needs and requirements in a series of previous studies (Bernhaupt *et al.* 2008, 2011).

Playful probing

Playful probing uses playful elements in the probing material including participatory design tasks that are based on modelling clay. To enhance the commitment of the participants in the study, a game is designed that addressed the thematic topic of the study. We typically use the standard set-up of cultural probing, taking for example postcards or post-its as probing material to gather insights on people's habits and media usage. The playful probing approach differs from the traditional approach as it uses games that are specially designed for the study. In playful probing, the games are designed to focus on the research area addressed within the study. The development of the game itself depends on the study set-up. Depending on the topic to

be investigated, variations of existing games can be used or even new games are developed. In opposition to other probing approaches, playful probing is a group activity involving at least two participants at the same time (Bernhaupt *et al.* 2011).

In order to gain a holistic understanding of people's practices and needs and the potential for new interaction techniques, we triangulated qualitative probing material (e.g. photos, videos, interview data, some of the creativity cards, sketches) and quantitative probes material (e.g. questionnaire data, some of the creativity cards). This does not go in line with the traditional cultural probing approach proposed by Gaver *et al.* (2004), who only used qualitative probes materials as source of inspiration for design concepts. However, the triangulation of various kinds of materials allowed us to be more confident of the collected data (consistency check) and it stimulated a richer understanding of household's technology usage patterns by addressing different aspects (Jick 1979).

Households and participants

The ethnographically oriented study using the playful probing method was conducted in two countries. The first part of the investigation was performed in France using a high-technology-oriented sample of IPTV users, while the second part was conducted in Austria with a sample of eight different user groups (selected by different media and entertainment consumption behaviours). The Austrian study has been conducted within a broader ethnographic study focusing also on other research topics which are not presented in this article.

French sample

The French part of the study was conducted in November 2011, recruiting 15 households in Toulouse, France, by local announcements, mailing lists and recruiting via a well-known major social network. Households did not know each other. In total, 30 participants older than eight years participated in the study. As one of the goals of the study was to identify different usage practices and needs related to the usability and the overall user experience of a smart phone application, three different groups of households were recruited: (a) households with (technologically oriented) singles, (b) households with (technologically oriented) young couples or families that finished their education and (c) households with families including at least one child older than eight years. Each group had five households.

All households had as primary TV with one of the available IPTV offers in France (eight had Free, three had the offer called Neuf, one Bouygues, three did not provide details about their IPTV offer). The IPTV offers all included triple play (TV, Internet and phone via IP). For 13 of the IPTV offers, there is already a smart phone application

available that allows to directly control the IPTV set-top box given that the mobile phone is in the same wireless local area network (LAN). The solutions do not support controlling any other device in the living room.

All households had broadband Internet access, 10 had activated the local wireless network from the IPTV provider, and seven additionally had a LAN. At least eight of the households had access to the Internet via the set-top box available. Households had between one and six remote controls in their living room to control the (entertainment-related) devices around TV, only one household owned a URC, which was not in use.

The main TV was positioned in the living room, five households had a second TV positioned in the bedroom, and one household had three TVs. All TVs in the households received digital TV, as analogue TV is no longer available in this region.

Adult participants in the household were between 19 and 51 years old ($n = 22$; 9 females and 13 males), with an average age of 33.27 (SD = 9.8). People obtained a broad range of education levels, which is typical for this region, and professions included secretary, hairdresser, sales person, technical engineer and household help. Some were students and one person was unemployed. The majority of participants (15) had French as their primary language, seven indicated to be bi-lingual (three Arabic, two Tamil, one Creole, one Chinese). All participants indicated to speak at least one additional language (English/Spanish). Choosing multilingual participants should allow gaining insights in needs related to accessing international offers and content that is typically available in various languages.

Austrian sample

Overall, 56 households participated in the study; 120 persons (68 female and 52 male persons), aged from 2 to 84 years, were living in these households, in which 94 persons (53 women and 41 men) were older than or equal to 18 years (18–84, $m = 40.66$, SD = 14.6) and 26 persons were younger than 18 years (26 children, 2–17 years, $m = 11.96$, SD = 4.56, 15 girls, 11 boys). The number of persons living in the households ranged from one to five persons, with 19 households consisting of three or more persons, 16 couples and 19 single households.

The households were selected based on a pre-specification of seven different media and interactivity groups: heavy media users (HMU), referring to people that use all kinds of media and consume media on a variety of devices; heavy TV users (HTVU), referring to people that use TV more than the average and that typically structure their lives with the help of TV; distributed media users (DMU), referring to people that use media not necessarily with different devices, but on different locations (e.g. in a second home); DINK/DIOKs, referring to double income no kids or double income one kid; families, representing typical families in Austria (specifically in the

'Middle-Class'); Silver Sofas, referring to people aged 55+ that dispose over more time than working people; and finally the so called sporadic users, who consume media on an irregular basis.

Households were recruited in three regions of Austria (Salzburg, Carinthia and Vienna), based on a multitude of criteria in terms of user group (media behaviour), but also type of TV reception(s), balancing of city vs countryside and balancing of various interests due to the geographical nature of the country (e.g. regions close to borders do have different interests in terms of media reception/infrastructure).

The (main) TV is physically present in the living room in all the households (95%, $n = 53$); the second TV (TV2) is typically in the bedroom (63%, $n = 17$) followed by the children's room (25.9%, $n = 7$). The third and fourth TVs (TV3 and TV4) are typically positioned in the bedroom, the kids' room or other rooms like guest rooms/working rooms. No TV was reported to be in a kitchen.

Among the participants, 20% were office employees, 19% in services, 14% in technical occupations, 7% in non-technical sales occupations, 8% white-collar employees and 24% were other employment types. All but one household were native German speakers.

Procedure and material

The interviewer visited the households in both countries at the beginning of the study and provided a questionnaire to focus on devices and equipment that were currently installed in the household and the playful probe package (see Figure 1 as an example) including creativity probing cards, modelling clay, pens and post-it notes. The French households received a specially designed game called '1000 miles' that focused on the central research questions on interaction techniques, applications and overall awareness building, while the game in Austria was a variation of a game called Personality – asking the participants to construct a TV product that best fits their needs (including questions related to the control of this TV product). The modelling clay was added to the probe package to encourage the participants to build some designs of simplified versions of the controls they would want to have.

Figure 1.   Example of playful probing material distributed to the households.

The participants were asked to document important aspects of their daily life by recording videos and taking photos as well as solving tasks on the creativity cards.

At the end of the one-week observation period in France and of the three-week observation period in Austria, the interviewer visited the households again to conduct a final interview. Final interviews were conducted (wherever possible) with all household members at the same time. In the final interview, the probing material was discussed with the participants using additional questions to focus especially on the experience of the households regarding interaction techniques that could be supported by standard smart phones today (touch interaction, gesture, voice).

Media usage in the French sample

Media usage is reported based on the quantitative probing material. From the 22 adult participants, 17 participants watched TV on a daily basis, five participants watched several times a week. Table 1 shows the indicated usages for various media types. The usage of media is rather high, which is in line with the three household types we recruited. In addition, 13 participants have been installing applications on their mobile phone, and 10 have been searching application for their mobile phone in the Internet. *None* of the participants *used the available application to control the set-top box* in their home.

Eight participants in the study were children between 8 and 17 years of age, mean 12.5 (SD = 3.5). They reported to have watched TV at least several times a week on average

Table 1. Media usage frequencies of the 22 adult participants.

Media	D	STW	STM	R	N	AVG (SD)
TV	17	5				129 (74)
Radio	6	5	1	3	6	44 (38)
CDs	2	4	3	1	8	45 (28)
Reading	9	3	3	4	2	64 (35)
Notebook/PC	16	1	2		1	172 (155)
Video games/PC	1	4	2	4	8	251 (243)
Video games/Console	2	3	2	7	7	82 (83)
Internet (private)	19	1	2			–
Facebook	7	2	1	4	7	–
Mobile phone (calls private)	15	5	1	1		–
Internet via mobile phone	9	4	1	0	6	–
E-mail via mobile phone	11	3	1	0	6	–
SMS	11	7	2	1		–
Photo/video mobile phone	2	8	2	9	1	–
Music mobile phone	4	3	1	1	7	–

Note: D, daily; STW, several times a week; STM, several times a month; R, rarely; N, never; AVG(SD), average time spent per day in minutes with this activity (standard deviation in min).

153 minutes per day (SD = 98). Especially the usage of mobile phones, Internet and Facebook was high among the children with the majority on a daily basis.

Thirteen households reported on the probing cards that TV (including the set-top box) is still the most often used device in the living room, one household mentioned the PC as most used, one household the X-Box (in France, there are various offers to reveice TV via the X-Box).

Media consumption in the Austrian sample

Media usage is reported based on the quantitative probing material. Media consumption in terms of watching TV, reading, listening to the radio is roughly similar to usage statistics on media consumption in Austria in 2011. The median of the TV usage is close to 2.5 hours per day (a median cuts the sample in half). This means that 50% of the participants of this sample (answering the questionnaire) watch TV for more than 2.5 hours per day, and 50% of the sample watch TV for more than 1.5 hours per day with others (social usage).

The median for private Internet usage is around one hour per day, with a very limited time using the Internet together. In general, it is clear that TV is still a social media (device), compared to other media and devices.

In the study, 60% of the participants indicated that they sometimes, rarely or never use an EPG, while 35% use it often or almost always and 5% of participants were unsure or did not reply. The EPG is not only used to retrieve information but also used to select channels (about 30% of the users using the EPG use it always or often to select a channel).

The number of mobile phones that are actually in use (with SIM card) ranges up to five mobile phones (including smart phones) in the study's sample. One household indicated to use five mobile phones (1.9%), seven households indicated to use four mobile phones (13.2%), 10 households indicated to use three mobile phones (18.9%), 11 households indicated to use two mobile phones (20.8%) and 19 households indicated to use one mobile phone (35.8%). Five households indicated not to use a mobile phone (9.4%, $n = 53$).

In the study, 60 participants indicated to own a smart phone, mentioning the following brands: 15 iPhone, 21 Samsung, 4 HTC, 2 Blackberry, 7 Nokia, 4 Sony Ericson, 3 LG, 4 other brands. As for the operating systems, 21 named Android, 9 iOS, 3 Windows, 1 Bada, and 26 indicated that they do not know the name of their operating system.

Regarding the use of apps, 39 participants indicated that they use apps on their smart phone, from which 9 use entertainment apps, 29 use social apps, 24 use game apps, 7 TV apps, and 14 personal apps.

Results

The results of the study were categorised using grounded theory (Strauss and Corbin 1990) as methodological

background. Results from France were categorised by three native speakers on an agreement basis, while results from Austria were categorised independently by two researchers with a correlation agreement of over 92%. Based on the first open coding, the axial coding focused on the research questions that guided this study. A variety of other dimensions including requests for enhanced media content, broader media offers, etc. were found, which are not reported in this article.

Direct control and navigation

Participants were asked on the probing cards to protocol what functions and buttons on the remote control were really important to them and what buttons they would keep for a simplified control. For IPTV households (French study sample), in terms of direct control all but one household agreed on the necessity of a simplified remote, either by reducing the number of buttons or by regrouping buttons. The most important functions were changing the TV channel, Volume control, and Navigations keys including ok. The households all mentioned the importance of changing the volume of a specific device (e.g. controlling the volume for DVD player, TV or set-top box). Three households mentioned explicitly pause/record, one household mentioned the colour keys (which are used to access important functions in the IPTV system the household used) and one household mentioned the need to access the teletext service on TV (to access language subtitles).

For the TV-oriented Austrian sample overall a reduction of keys on the remote control was of central importance (mentioned by 27 households) followed by enhanced usability (20 households). The problem of volume control was not mentioned in this sample even if households with satellite reception do have the same problem of needing to control TV and the set-top box. These statements replicate findings from a recent study where the buttons used on IR-based remote controls in households were logged (Mirlacher and Bernhaupt 2011). The study identified the numeric key-pad, navigation keys, volume keys, mute and on/off as the mainly used keys on standard IR-based remote controls.

We prompted participants from the French sample to think about how a smart phone application can help to overcome the identified problems:

(1) Remote controls typically offer direct access to a range of services. Participants reported usability problems with labels that were difficult to understand and remember, or simply meaningless to them.

> I do not understand this abbreviation, the [smart phone] app should just have a longer label, that would make it clearer to me.[3] (France, household 2, female adult)
> I just don't know what happens when you press this button, there should be written something else. (France, Household 3)

To overcome the identified problems and limitations like the need for a reduced set of keys or usability problems with button labels, as well as to incorporate findings from both sets of the most important keys, named by the test participants and identified via IR-data logging, the following demands on smart phone applications are proposed: a smart phone application should provide direct control of IPTV and entertainment-related devices including direct control of channels and volume, basic navigation, and a limited number of usage-oriented scenarios (e.g. turn all devices on/off).

(2) A regrouping of the buttons on the smart phone application can support users to perform common tasks in the living room in a more efficient way. Additionally, a regrouping of buttons offers the opportunity to adapt the layout of the remote control to the individual user needs.

(3) Minimal design: study participants reported that they observed that they only need a minimal set of buttons to control their iTV experience, while a majority of buttons on the remote controls is rarely or never used: the application should therefore simply offer a minimal set of buttons, and all other functions should be made available in the interface displayed on TV.

(4) Simplification of long tasks: to support users that did not manage to solve complex tasks (e.g. accessing the latest recordings on the external hard drive), the application shall allow simplifying these tasks by providing a one-button solution (i.e. predefined or customisable macro buttons, or shortcuts to functions that need many steps to reach them).

(5) Flexible display of (touch based) buttons: the application should just present buttons that are related to the currently performed task.

(6) Participants stated that the mobile phone itself would help to overcome some recurring usability problems, like, for example, the missing backlight on some of the remote controls, making it impossible for them to use the remote controls in the dark. They mentioned that the application should have the ability to reboot devices (mobile phone app is informing about the need to update the device via a reboot). Other needs were to enhance the switching between channels which is perceived as slow.

In the Austrian standard TV sample, the majority of users stated that they were not interested in a smart phone application overall, and mobile phone applications for TV are currently not perceived as an integral part of the future TV experience. The control of entertainment devices was only mentioned by one household who was designing the TV product of their dreams including a tablet-touch control
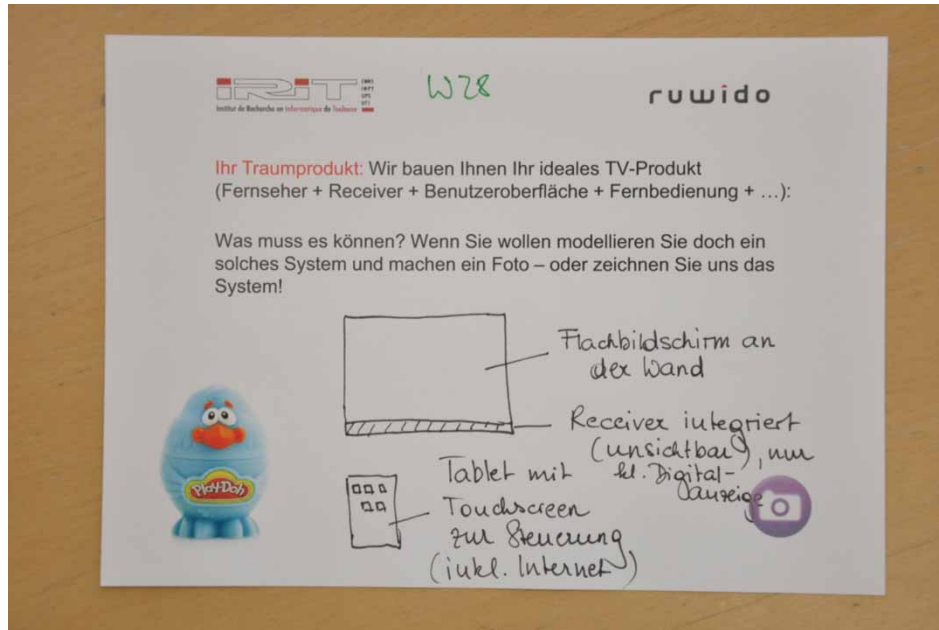
Figure 2.   Depiction of the TV product of their dreams including a tablet with touch-control.

(Figure 2), and in general, participants still use alternative devices or media for any additional activity they want to perform ('I look it up on Internet'). Overall the standard TV sample was not interested in additional interactivity (31 out of 56 households) with seven participants mentioning an interest in a low level of interactivity, four medium interactivity and five high interactivity (user groups: heavy media users, heavy TV users, families).

The majority of participants were not appreciating the idea of using a tablet PC or smart phone to control the functionalities of TV or getting additional information. The heavy media user group was the least negative and had by far the highest value for a positive opinion towards this concept (50%), while families, Silver Sofa and distributed media user groups had the largest value with more than 70% of each group against this concept.

Being asked whether they would prefer to have apps on their TV, more than half of the participants feel reluctant to using apps in their TV sets – they do not see the need for apps, are afraid they cannot handle them, or prefer using apps on a mobile phone.

Interestingly, participants nevertheless have a quite clear idea which apps they would want to use on a mobile phone, on a TV, or on both screens.

Enhancement of interaction technique

IPTV households (France)

Touch was the preferred type of interaction with the mobile phone, but was also seen as a preferable means to enhance the current forms of remote controls. Touch was mentioned as being helpful for personalisation of the controls:

I simply want to have a list of buttons and then I drag the important ones on one page. (France, Household 5)

When applying touch to any form of remote control, users requested to have a combination of both buttons and a touchpad on the remote control to perform the most important tasks. A participant suggested having a touchpad to write down the three numbers with the finger in order to change to a specific channel:

I know that 637 is my favourite foreign language channel, and I could simply write that down. (France, Household 7)

Gestures were less welcomed by the users. Due to the limited functionality of the gesture-based remote control that is currently in their households, users neglected to use gestures:

I tried to navigate the Internet with these gestures, but I immediately gave up and went to my PC. (France, Household 8)

Ideas on gesture usage were more related to enabling to point with the mobile phone towards a device or user interface element. 'I just point on the DVD [player] and then control the volume directly on it.' French household 2 wanted the same type of selection of a device by pointing at it, but simply to change the language of the programme instead of changing the volume. One household member (France, Household 12) suggested enabling interaction with TV like in 'minority report, you know, where everything is see-through and you can just move objects around'.

To enhance direct access to frequently used services or functions, users described speech input as a possible means. The speech input should consist only of a limited set of commands, which 'you can speak directly into the mobile

phone' (France, Household 9). In Household 2, one of the children (girl, 15) made a design including speech as one of the icons that can be selected in the main menu of TV, enabling speech input.

To improve the interaction with the IPTV system, participants requested to have only one device in their hands that allows them a simple control, and the IPTV system would be intelligent enough to know what they would want to do or watch.

> I just want to have one button, with a good tactile feedback, and the menu on TV is simply showing me the things I want to get.

Standard Austrian sample

While there was low interest in control via smart phone, the majority of households argued for a dedicated device (a 'new version of a remote') that might incorporate some additional elements like wheels or small touch areas. The Austrian sample wanted a device that is devoted to the control of TV and helps them overcome their usability problems.

Usage-oriented scenarios (IPTV households, France)

To enhance the overall user experience, IPTV households suggest having various functions, tasks and activities grouped together. The needs and requirements stated were broad, what they had in common is that they were all around certain usage situations of the entertainment system or IPTV system. We refer to them as usage-oriented scenarios.

Supporting the user experience of watching TV

IPTV-participants reported to watch TV still on a daily basis or several times a week. A smart phone application would best support their TV experience enabling them to watch TV: simply changing channels and changing the volume are some of the basic requirements. While this sounds obvious as a request, changing volume has become a complicated task involving the volume control of the set-top box and the volume control of TV. Typically two remote controls have to be used to set the volume to an agreeable level; furthermore, the user has to take care that the sound level of TV and the set-top box is balanced, as a too high volume setting on TV may result in a moment of shock when changing the input source and the audio signal is very loud. Also the procedure to simply watch TV is in some cases quite cumbersome and requires two remote controls.

(1) It is necessary to select the input or source the signal is coming from (typically labelled 'source' on the remote control), as this allows switching between the various connected devices including the set-top box.

(2) Once the set-top box is selected, the TV offer has to be chosen in the user interface of the set-top box, for example the standard offer of the IPTV provider or the premium content (using the remote control of the set- top box).

(3) Then the user can switch between channels.

Households mentioned additionally that the desired volume level is depending on the time of the day:

> You know, when I come home it is just running in the background, but when I watch a movie after dinner I want automatically a different volume value. (France, Household 11, male, 27)

To enhance the user experience of watching TV, a smart phone application should allow the user to simply control the volume and access the TV mode in an easy way, hiding any necessary adjustment of different volume regulations on different devices from the user and allowing access to TV without requiring any menu selection or other activities.

Usage-oriented control

Beyond the need to enable the simple activity of watching TV, users stated that their most often used activities include watching DVDs, listening to music or playing games. To enhance the control of such usage-oriented scenarios, the users suggested having a set of buttons on one page of the mobile phone application, allowing them to control these activities. The application should additionally allow to set-up, modify and delete such control settings that are related to usage-oriented scenarios. The usage-oriented control (households called them 'control pages') should be easily accessible in one menu and should adapt to the current usage situation or personal preferences chosen.

Turning on and off

Enjoying any type of moving pictures in a living room means to turn on and off a multitude of devices. Depending on the usage-scenario (e.g. watching TV, watching a DVD or simply listening to music) this can involve various different devices. Users expressed their wish to have an intelligent function that automatically turns on and off the requested devices for the various usage scenarios. Other households suggested to have just all devices turned on and off, independent of the usage scenario.

The mobile phone application could support this user needs by offering shortcuts or one-button solutions with predefined functionalities (e.g. turn on TV and IPTV box) or customisable functionalities (e.g. turn on radio, sound system and amplifier).

Selecting and handling content

Naturally, the first step to enjoy any type of content is to be able to choose and select it. Content selection has become a challenge: users reported that they have a multitude of devices or what they called 'places' to look for content. This includes for example the hard disk of the set-top box for recorded programmes, movies that were bought on different video on demand (VOD) portals (in the IPTV systems

the household owned, there is a range on VOD offers available, and access to (rented) movies requires to navigate to the appropriate portal), content on DVDs, content on external (additional) hard disks, content that is accessed via the local area network, content that is streamed from the Internet (e.g. YouTube), content that is accessed via the game consoles (either reading a DVD, receiving premium content or accessing content via the Internet) and content that was downloaded on a PC and transferred to TV via a USB stick.

The smart phone application should enable them to display all possible 'places' where content can be accessed from a menu providing a general (device independent) overview. This would allow them to easily select the content, without the need to navigate through several interfaces and without the need to change (or even rewire) devices.

A users' wish that is closely related to supporting such unified views is to display all connected devices. This primarily serves the purpose as a first step to select the content source, but also as a means to understand what devices are currently actively connected.

> You know, I am really upset when I am already installed on the couch and then I realize that the kids unplugged the [DVD] player, to be able to connect their game [game console]. (France, Household 3)

Service usages including the home context

Users presented a set of scenarios that can be supported by the smart phone application, including control of the PC, telephone and lights, as well as personal calendar or access to information on personal devices and services (e.g. reminders, …). Especially for the control of web pages on the TV screen they suggest to enhance the interaction with number and text input on smart phone.

A notable difference between the two samples in Austria and France was the usage of teletext, which is far more common in Austria. During the study, participants in Austria were asked which features they would like to have for the TV of the future. Nineteen household would welcome features like news, weather information, etc., similar to a smart phone. Nineteen households do not want this information, and further seven households still rely on teletext to receive this kind of information. Another four households prefer to use TV to get this information, two households would want it for their children only ('to be connected') and one household has different opinions among the family members ($n = 52$). These functions are most appreciated by heavy TV users and sporadic users (four out of eight households each) and least wanted by the Silver Sofa group, which is also relying the most on teletext.

Interestingly, answers were quite contrary when the participants were asked which applications they would like to have the most on their TVs. Overall, 49 households answered this question and named 63 applications:

- 30.2% of answers stated that that the households do not want applications on their TV (mentioned most often by sporadic, Silver Sofa and heavy media user groups);
- 22.2% mentioned Internet, Wikipedia, Skype, Internetsuche, Google-Earth, YouTube (mentioned most often by DMU, HTVU and DINK user groups);
- 9.5% named news (available anytime), information about weather and sports, Lotto & Toto (betting) (most often named by sporadic and families user groups);
- 9.5% mentioned games (mentioned most often by DINK user group);
- 6.4% mentioned integrated CD, DVD, VCR (most often by families, followed by DINK and Silver Sofa user groups);
- 6.4% mentioned mobile phones and TV: SMS, calls, video calls, contacts, calendar (mentioned most often by Silver Sofa user group);
- 4.76% photo library, watching photographs (most often by DMU user group)
- Radio and music was mentioned in 3.17% of answers (HMU, HTVU user groups).

The following were named only one time each (1.59% each):

- should be the same as on the iPhone: 1.59% (sporadic user group);
- to control the home automation from the living room: 1.59% (HMU user group);
- to record broadcasts: 1.59% (DMU user group);
- to watch new cinema movies: 1.59% (DMU user group);
- voice recognition: 1.59% (DMU user group).

Design ideas and two-screen design

The probing package requested household members to provide some design ideas on how to display information on the mobile phone and on the TV screen. Users primarily kept the current form of interaction: the TV screen is the user interface, displaying the content, while the mobile phone screen is simply used to control the information displayed on the TV screen. Information on the TV screen should be shown at the borders of the screen (not using the whole screen like the majority of current IPTV user interfaces). The more important the information is becoming, the more space of the screen should be used.

This confirms industry attempts to design the next generation of user interfaces following the level of engagement of the users: if users are engaged only to a limited degree (e.g. they just choose the main menu), this should not hinder them from enjoying the (TV) content, but the more selections the user is making (e.g. the user chooses a sub-menu or additional information), the more space can be used.

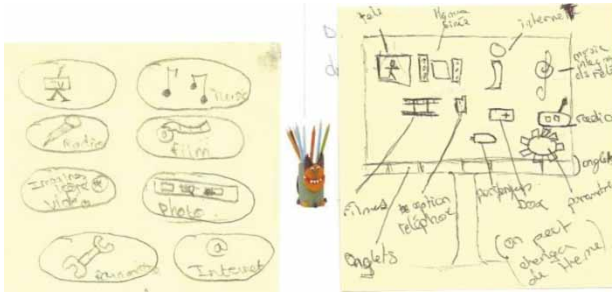Designs from the users ranged from iconic views ('they simply should use the same icon view as I have on the

Figure 3. Depiction of two designs that were suggested by Household 2, France (girl, 17, girl, 13.5). Noticeable is that both drawing suggest Internet as important category (left: button down to the right; right: depiction 'i'.

iPhone', France, Household 2, girl, 17 years.), to presenting the information as a pie chart or as a list (France, Household 2, mother; France, Household 2, girl, 15.5 years). Figure 3 shows some of these designs. The majority of households referred to the 'standard design' currently available (a bar, enabling the user to select items navigating left or right), or a mosaic view. Users that were speaking Arabic/Tamil preferred the user interface to be positioned at the right side of the screen.

Households today are comfortable with the idea that the user interface of TV is structured similar to other devices like PC, mobile phone or web pages. Household 15 (France) suggested to have a Windows-oriented graphical user interface, to enable listening to the music at the same times as selecting information in the user interface displayed on TV.

IPTV users' differing needs for usability and user experience of control applications

User practices, needs and requirements were different in the three IPTV household groups. Table 2 shows an overview on households' needs in terms of usability improvements and households' suggestions to improve the overall user experience.

In terms of privacy, the participants saw the mobile phone as an ideal means to achieve privacy: private information should be displayed on the phone while information (for example photos) should be shared on the TV screen. The mobile phone was indicated to be helpful for personalisation [F-HH6: 'I can get a list of all the movies I like on the mobile phone, and I do not have to look that up on the Internet again (referring to using Internet on the PC)']. Smart phone could support people in terms of recording: either by suggestions on what to record on mobile phone or by helping to administer (one central) list of recordings for all entertainment devices. To enhance the overall user experience, participants suggested to enable access to social networks on TV or in combination with mobile phone, and to display additional information that accompanies the running programme.

Table 2. User needs mentioned in the study related to usability and user experience in the French study sample, split by household groups.

	Households with children	Technology oriented, single	Technology oriented, couple
Needs/requirements related to the usage of the system (usability)			
Less functions; less usability problems with buttons	4	4	2
Less actions necessary, faster actions	4	5	3
Easier learning	3	4	3
Less memorising (less cognitive load)	3	4	3
URC	1	1	2
Context awareness	2	5	1
Central control	1	1	2
Help to solve problems	1	2	1
Enhance feedback	1	0	2
Lights	0	0	1
Needs/requirements to enhance the overall user experience			
Privacy	3	5	3
Profiles/ personalisation	1	2	2
Easier recording of programmes	4	2	4
Social networks	2	3	3
Additional/enhanced information to current programme	1	2	0

Parallel usages of media (Austrian sample)

Given the above insights from the French study it was rather surprising that overall the standard Austrian TV user sample was less interested in controlling their environment via smart phone applications. This might have various explanations, but given such an application is available it is interesting to know if users already use other devices while watching TV. To further investigate parallel usage of other devices and media, the Austrian participants were asked how often they use various devices in parallel to watching TV. The results are presented in Table 3.

Given the high number of people using their mobile phones while watching TV, it is clear that mobile phone is available and thus might also be used for controlling TV (Table 3).

Design recommendations

Based on the user requirements identified in the field study using a playful probing approach, the following design recommendations have been found helpful for the development of a smart phone application.

(1) Enable the control of all devices related to the IPTV experience. Current smart phone applications limit

Table 3. Parallel usages of other devices and media while watching TV in the Austrian sample. Users in the Austrian sample are using a variety of media while watching TV.

Watching TV and using	Always	Often	Sometimes	Rarely	Never
Tablet	1	2	4	2	83
Laptop	4	17	19	12	40
Reading (newspaper, journal, …)	1	17	21	18	35
Internet on PC/laptop	6	17	13	10	46
Video games	2	3	4	1	82
Making calls on the phone	6	17	30	20	19

the user to the control of one device (e.g. the set-top box), to a set of devices that have one type of communication protocol (e.g. all infrared controllable devices) or to the control of all devices that are in the same wireless LAN. From a user's perspective this means that the smart phone application is just another form of remote control, with the need to use additional remote controls for various tasks. To enable users to overcome current usability problems when accessing media from a set of entertainment devices, the application must allow to control all the devices in the living room. This includes the support of controlling devices via infrared (IR), various radio frequencies (RF, RF4CE, Bluetooth, Bluetooth LE) and wireless LAN. This might include the necessity to provide additional infrastructure that goes beyond standard solutions.

(2) Design for 'usage-oriented' scenarios. When designing the application it is important to help the user perform the most important tasks, independent of the number of devices involved. This requires a detailed analysis of the tasks users typically perform with their IPTV system, including watching TV, buying and watching movies, watching DVDs, or playing music. The mobile phone application should allow the user to select one of these activities and provide the most important controls related to that activity. If possible the control should be supported by some intelligent mechanisms (e.g. in the case of automatic turn on/off of the devices, or when information shall be accessed across devices).

(3) Design for personalisation and personal usage. To support the user in their personal usages, the interface must enable the user to flexibly generate 'quick access pages' for the most often used controls, devices, functions and services. The ability to perform a sequence of control events (for example as macros) should be included. Lessons learned from URCs clearly indicate that macros are difficult for users to develop, maintain and programme. We thus recommend supporting the user with a set of predefined macros or usage scenarios that can be adapted to fit the personal needs and the technological infrastructure in the households.

(4) Support the user to control the connected home. The control of any device in the household can be a convincing argument for the continuous usage of smart phone based control applications. Especially the integration of other devices can simplify the control overall and provide thus the necessary meaning and value for the acceptance of such applications.

(5) Enhance the overall user experience by providing meaningful functionalities that satisfy users' needs. A smart phone application can be helpful to enhance the overall experience. Changing a channel today can need up to three seconds. People do enjoy the experience to 'see what is on other channels', but with channel switching times of up to three seconds this becomes tedious. Smart phone applications can bring back this experience, allowing the user to flip through lists of channels and their content description, even allowing the user to preview selected content. By offering such a feature, the usability problem of slow channel switching can be transformed into an enjoyable and positive user experience.

Other possibilities to enhance the user experience include enabling the support of privacy and security. While a TV today is still used by several members of the household, mobile phone typically is a personal device. The smart phone app thus can be used to control access to special content (e.g. payment can be done via mobile phone), as well as to allow the user to control what is displayed on the large screen (e.g. what photos shall be shown on the TV screen, what should stay private). Displaying content on the large screen can additionally include context dependent display of services, where e.g. the user can select on smart phone if she wants to have the latest twitter messages displayed on the big screen or if she prefers this service to be only shown on the private device.

(6) Support touch and speech as interaction techniques. Users are open to new forms of controls in their living room. Touch interaction with a smart phone has become widely accepted and was also suggested as an enhancement of current standard remote controls. Gesture was less welcome due to negative

experiences with the currently available systems, but was seen as important to select devices or items by a pointing gesture. Using a limited set of speech commands is a promising approach, due to the ability of a smart phone processor to process language (maybe by simply processing speech recognition in the cloud).

(7) Be aware of limitations: consider your audience. Given the results from the two samples it becomes obvious that a mobile phone application for the control of entertainment devices will not be used by all user groups and audiences. The mobile phone control app can be helpful to overcome usability problems in IPTV environments as well as in other TV environments. Given that a high number of participants indicated to make calls while watching TV, the mobile phone might not be the best choice for such a control application. What is important is to take into account that such applications will currently never reach all user groups, and thus an alternative (like the standard remote) for the control of the entertainment environment is still necessary.

Summary and discussion

This study presents the results from a field study using the playful probing approach to investigate user needs and requirements related to smart phone applications that enable the control of all entertainment devices. While the majority of insights seem to simply confirm our daily experiences with entertainment devices, insights in this study are unique as participating households owned the latest generation of IPTV systems including Internet access, remote controls that supported gesture and the ability to use smart phone applications that allow the control of the set-top box.

To summarise, a smart phone application that aims to be successful in terms of usability and user experience should support the user following these key recommendations: (a) allow direct control of all devices in the living room, (b) design for and support usage-oriented scenarios, (c) enable personalisation and personal usage, (d) provide meaningful functionalities that satisfy user needs to improve the overall user experience, (e) support touch and speech as interaction modalities and (f) consider your audience and be aware of limitations. Additionally any smart phone application can be helpful to support users in their daily usage and to overcome existing usability problems (Table 2).

What seems to be clear is that the current generation of smart phone applications is likely to fail, as it does not support users in their initial wish to have only one central control for all their entertainment devices.

In terms of interaction techniques, the usage of smart phone applications is a promising step, as flexible touch interaction, speech and a limited set of gestures can be used. This is partly in-line with recent findings in the literature that touch has become an acceptable means of interaction.

It replicates and supports findings from Kela *et al.* (2006) that gesture interaction has to be carefully designed in terms of the number and type of gestures and that the simple button remote control might outperform gestures (Bailly *et al.* 2011a). Using a smart phone for interaction via speech seems acceptable for the users. From a user perspective, talking to a smart phone might feel almost natural, compared to talking to a remote control.

The development of an application taking into account the above design recommendations will be a complicated endeavour. It involves the development of new types of hardware (to control IR, RF, WLAN) and new forms of (bi-directional) connections (to detect all devices that are connected to a TV or in the entertainment set-up), and might require changes in current software on almost all devices in the living room. We are currently developing a set of new applications that will be based on the above design recommendations.

Notes

1. http://www.apple.com/uk/apps/remote/
2. http://l5remote.com/, http://www.blinqtv.com/
3. All statements were translated from French.

References

Bailly, G., *et al.*, 2011a. Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. *Proceedings of Interact 2011*, LNCS 6947, 248–262. doi:10.1007/978-3-642-23771-3–19

Bailly, G., *et al.*, 2011b. Gesture-aware remote controls: guidelines and interaction techniques. *In*: *Proceedings of ICME 2001*, 263–270. doi:10.1145/2070481.2070530

Ballagas, R., *et al.*, 2006. The smart phone: A ubiquitous input device. *IEEE Pervasive Computing*, 5 (1), 70–77.

Bernhaupt, R., *et al.*, 2007. Playful probing: making probing more fun. *INTERACT 2007*, 10–14 September 2007. Springer LNCS, 606–619.

Bernhaupt, R., *et al.*, 2008. Trends in the living room and beyond: results from ethnographic studies using creative and playful probing. *Computers in Entertainment*, 6 (1), Article 5 (May 2008), 23. doi:10.1145/1350843.1350848

Bernhaupt, R., *et al.*, 2011. Security, privacy, and personalisation: informing next-generation interaction concepts for interactive TV systems. *Computers in Entertainment*, 9 (3), Article 17 (November 2011), 33. doi:10.1145/2027456.2027463

Cesar, P., Chorianopoulos, K., and Jensen, J.F., 2008. Social television and user interaction. *Computers in Entertainment*, 6 (1) (May 2008), 1–10. doi:10.1145/1350843.1350847

Darnell, M.J., 2008. Making digital TV easier for less-technically-inclined people. *In*: *Proceedings of the First international conference on designing interactive user experiences for TV and video*. Silicon Valley, CA, USA, October 22–24, 2008. UXTV '08, vol. 291.

Dezfuli, N., *et al.*, 2012. Leveraging the palm surface as an eyes-free TV remote control. *In*: *CHI'12 extended abstracts on human factors in computing systems*. New York: ACM, 2483–2488.

Elder, H., 1970. On the feasibility of voice input to an online computer processing system. *Communications of the ACM*, 13 (6), 339–346.

Enns, N.R.N. and MacKenzie, I.S., 1998. Touchpad-based remote control devices. *In*: *CHI 98 Conference summary on human factors in computing systems (CHI '98)*. New York: ACM, 229–230. doi:10.1145/286498.286717

Epelde, G., *et al.*, 2009. URC based accessible TV. *In*: *Proceedings of the seventh European conference on European interactive television conference (EuroITV '09)*. New York: ACM, 111–114. doi:10.1145/1542084.1542103

Furnas, G.W., *et al.*, 1987. The vocabulary problem in human-system communication. *Communications of the ACM*, 30 (11), 964–971.

France Mediametrie, 2012. *L'audience de la télévision en 2012* [online]. http://www.mediametrie.fr/television/communiques/l-audience-de-la-television-en-2012.php?id=789#.UYfD-0qnDg0 [Accessed 15 May 2013].

Gauntlett, D. and Hill, A., 1999. *TV living: television, culture and everyday life*. London: Routledge.

Gaver, W.W., *et al.*, 2004. Cultural probes and the value of uncertainty. *Interactions*, 11 (5) (September 2004), 53–56. doi:10.1145/1015530.1015555

Gawlinski, M., 2003. *Interactive television production*. Oxford: Focal Press.

Goldberg, D. and Richardson, C., 1993. Touch-typing with a stylus. *In*: S. Ashlund, A. Henderson, E. Hollnagel, K. Mullet, and T. White, eds. *Proceedings of the INTERCHI '93 conference on human factors in computing systems (INTERCHI '93)*. Amsterdam, The Netherlands: IOS Press, 80–87.

Goto, J., *et al.*, 2003. A spoken dialogue interface for TV operations based on data collected by using WOZ method. *In*: *Proceedings of the 41st annual meeting on Association for Computational Linguistics*, 2003, 101–104. doi:10.3115/1075178.1075191

Harrison, C., Benko, H., and Wilson, A.D. 2011. OmniTouch: wearable multitouch interaction everywhere. *In*: *Proceedings of the 24th annual ACM symposium on user interface software and technology (UIST '11)*. New York: ACM, 441–450. doi:10.1145/2047196.2047255

Hess, J., *et al.*, 2011. Using paper and pen to control home-IT: lessons learned by hands-on experience. *In*: *Proceedings of the 9th international interactive conference on interactive television*. New York: ACM, 203–212.

Igarashi, T. and Hinckley, K., 2000. Speed-dependent automatic zooming for browsing large documents. *In*: *Proceedings of the 13th annual ACM symposium on User interface software and technology (UIST '00)*. New York: ACM, 139–148. doi:10.1145/354401.354435

Igarashi, T. and Hughes, J.F., 2001. Voice as sound: using nonverbal voice input for interactive control. *In*: *Proceedings of the 14th annual ACM symposium on User interface software and technology (UIST '01)*. New York: ACM, 155–156. doi:10.1145/502348.502372

Jick, L., 1979. Mixing quantitative and qualitative methods: triangulation in action. *Administrative Science Quarterly*, 24 (4), 602–624.

Kela, J., *et al.*, 2006. Accelerometer-based gesture control for a design environment. *Personal Ubiquitous Computing* 10 (5) (July 2006), 285–299. doi:10.1007/s00779-005-0033-8

Lin, J., *et al.*, 2010. Free hand interface for controlling applications based on Wii remote IR sensor. *In*: *Proceedings of the 9th ACM SIGGRAPH conference on virtual-reality continuum and its applications in industry (VRCAI '10)*. New York: ACM, 139–142. doi:10.1145/1900179.1900207

Logan, R.J., *et al.*, 1995. Living room culture – an anthropological study of television usage behaviors. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 39(5), 326–330.

Lorenz, A. and Jentsch, M., 2010. The ambient media player – A media application remotely operated by the use of mobile devices and gestures. *In*: *Conference proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia (MUM '10)*. New York: ACM, 372–380. doi: 10.1145/1899475.1899490

Mirlacher, T. and Bernhaupt, R., 2011. Breaking myths: inferring interaction from infrared signals. *In*: *Proceedings of the 9th international interactive conference on Interactive television (EuroITV '11)*. New York: ACM, 137–140. doi:10.1145/2000119.2000146

O'Brien, J., *et al.*, 1999. At home with the technology: an in-home computing and ethnographic study of a set-top-box trial. *ACM Trans. Computer-Human Interaction*, 3 (3), 282–303.

Ofcom, 2007. *Easy to use digital television receivers: remote control buttons and functions used by different types of consumer* [online]. http://www.ofcom.org.uk/research/tv/reports/dso.pdf [Accessed 15 May 2013].

ORF Media Research Austria, 2012a. *ORF Media Research* [online]. http://mediaresearch.orf.at/c_fernsehen/console/console.htm?y=1&z=2 [Accessed 15 May 2013].

ORF Media Research Austria, 2012b. *ORF Media Research* [online]. http://mediaresearch.orf.at/index2.htm?fernsehen/fernsehen_heimel.htm [Accessed 15 May 2013].

Pelling, C., Sko, T., and Gardner. H.J., 2009. Be careful how you point that thing: Wiimote aiming for large displays. *In*: *Proceedings of the 21st annual conference of the Australian computer–human interaction special interest group: design: open 24/7 (OZCHI '09)*. New York: ACM, 397–400. doi:10.1145/1738826.1738908

Pirker, M., Bernhaupt, R., and Mirlacher, T., 2010. Investigating usability and user experience as possible entry barriers for touch interaction in the living room. *In*: *Proceedings of the 8th international interactive conference on interactive TV & Video (EuroITV '10)*. New York: ACM, 145–154. doi:10.1145/1809777.1809808

Robertson, S., *et al.*, 1996. Dual device user interface design: PDAs and interactive television. *In*: M.J. Tauber, ed. *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '96)*. New York: ACM, 79–86. doi:10.1145/238386.238408

Seager, W., Knoche, H., and Sasse, M., 2007. TV-centricity – requirements gathering for triple play services. *In*: *In Interactive TV: A Shared Experience - TICSP Adjunct Proceedings of EuroITV 2007*. TICSP Series, vol. 35, 274–278.

Stifelman, L., Elman, A., and Sullivan, A., 2013. Designing natural speech interactions for the living room. *In*: *CHI '13. Extended abstracts on human factors in computing systems (CHI EA '13)*. New York: ACM, 1215–1220. doi:10.1145/2468356.2468574

Strauss, A. and Corbin, J., 1990. *Basics of qualitative research: grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.

Tsekleves, E., *et al.*, 2007. Interacting with digital media at home via a second screen. *In*: *Ninth IEEE international symposium on multimedia – Workshops (ISMW 2007)*. Los Alamitos, CA: IEEE Computer Society, 201–206. doi: 10.1109/ISM.Workshops.2007.42 , 201–206.

Turunen, M., *et al.*, 2009. User expectations and user experience with different modalities in a mobile phone controlled home entertainment system. *In*: *Proceedings of the 11th international conference on human-computer interaction with mobile devices and services (MobileHCI '09)*. New York: ACM, Article 31, 4. doi:10.1145/1613858.1613898