A Preliminary Usability and Universal Design Evaluation of a Television App User Interface

Bikram AWALE, Pietro MURANO
OsloMet – Oslo Metropolitan University, Postboks 4, St. Olavs plass 0130, Oslo, Norway
s320806@oslomet.no, pietro.murano@oslomet.no

Abstract. This paper details a two-part evaluation of the Apple TV user interface. A Heuristic Evaluation and an evaluation using the seven universal design principles were carried out. The overall results showed that the Apple TV user interface has various failings in the design for users of diverse kinds. Suggestions for ways forward include designers using more rigorously the ISO standards and well established user interface design guidelines.

Keywords: Television user interface, Usability, Human Computer Interaction, Universal Design, Heuristic evaluation, ISO standards.

1. Introduction

The use of television (TV) apps which stream content and programmes via the internet have become popular in recent years. Technology giants have introduced their own TV streaming devices and applications to the market over the years. Google has created Chromecast (Google Inc, 2019), Amazon has created Fire TV Stick (Amazon.com, 2019) and Apple has developed Apple TV (Apple Inc, 2020a) for streaming digital content on TV via the internet. Apple TV allows consumers to use High Definition Television (HDTV) to stream video, music, podcasts, apps and games from Apple’s App store.

Most of these apps require some kind of interaction between the user and the system, where visual interaction features. To our knowledge, there has not been a systematic and detailed enough research investigation into the usability and universal design of such interactions with TV streaming systems and software.

This paper aims to begin leading the way in fostering better usability and universal design of such systems. We begin by detailing a preliminary evaluation of the Apple TV user interface as seen usually on a TV screen and operated with the Siri remote control. The evaluation will primarily concentrate on the on screen interaction and not specifically on the Siri remote control. However, the Siri remote control is mentioned below as interaction took place with this device. The authors would recommend that the Siri remote control should feature in a separate full evaluation as it has a series of specific design features that should be evaluated separately.
Apple TV was selected as the first ‘system’ we wanted to evaluate due to Apple having generally a high reputation for usability, user experience and a very loyal customer base. Many magazine type web sites reflect this viewpoint in their articles with at times one-sided sounding text, e.g., (Johnson, 2019). However, Apple is also quite unique in having some well known user interface experts who worked for Apple in the past, stating openly that generally Apple seems to have lost its way with design that is actually good for users (Norman and Tognazzini, 2015). Apple is also different in being a software/hardware company that has exceeded one billion users (Cybart, 2019) whilst controlling both hardware and software. Apple also supplies and controls an enormous ecosystem, which amongst many things allows one to access Apple TV via other Apple devices and certain other non-Apple devices (Cybart, 2019, Apple Inc, 2020c). Therefore, it seemed fitting to begin with an evaluation of Apple TV and in the future the prospect of continuing the research by also evaluating other well known media streamers by Google, Amazon and others would be a logical follow-on to this research.

Therefore, this paper will begin with a brief related literature section, followed by details of a Heuristic Evaluation and evaluation with the seven universal design principles. The paper will then conclude with a discussion and ways forward.

2. Brief Literature Review

Work on television user interfaces and apps for streaming television programmes has been done in the past. However, the authors of this paper upon examining the major literature in this context conclude that the amount of effort published on the usability and universal design of such user interfaces is rather limited. We therefore feel that more work needs to be done to address the usability and universal design issues that arise with such user interfaces.

Research by Chorianopoulos (2008) suggests that some developers and researchers have tried to address the issue of providing better systems for television interaction by devising design guidelines. Chorianopoulos (2008) devised some design principles for television applications. Chorianopoulos (2008) suggests that the design principles help with the development of television prototypes. He also suggests that these principles are a good help in maintaining a ‘trail’ of design decisions and rationales. Although the principles are a good idea, as Chorianopoulos (2008) acknowledges, they are currently rather wide ranging and would need to be more detailed to be more useful.

In a different study by Eronen and Vuorimaa (2000) the researchers developed two television user interface prototypes for digital television. One was meant to highlight simplicity while the other was meant to highlight efficiency. These were then tested with real users carrying some basic tasks. The researchers were not able to categorically conclude if one prototype was better than the other in terms of usability. However, some reasons for lack of conclusive evidence could be that their sample size was too small and the evaluation method used was perhaps not rigorous enough.

In a different study by Eronen and Vuorimaa (2000) the researchers developed two television user interface prototypes for digital television. One was meant to highlight simplicity while the other was meant to highlight efficiency. These were then tested with real users carrying some basic tasks. The researchers were not able to categorically conclude if one prototype was better than the other in terms of usability. However, some reasons for lack of conclusive evidence could be that their sample size was too small and the evaluation method used was perhaps not rigorous enough.

In a different study by Eronen and Vuorimaa (2000) the researchers developed two television user interface prototypes for digital television. One was meant to highlight simplicity while the other was meant to highlight efficiency. These were then tested with real users carrying some basic tasks. The researchers were not able to categorically conclude if one prototype was better than the other in terms of usability. However, some reasons for lack of conclusive evidence could be that their sample size was too small and the evaluation method used was perhaps not rigorous enough.

Acknowledging issues surrounding ease of use of televisions, Freeman and Weissman (1995) developed a way to navigate or interact with television using hand gestures. Their aim was to address the building of systems that are easy to use. Although limited information is given regarding the evaluation of the hand gestures, the authors suggest that more work would need to be done to find out if their approach would be good enough in a user’s home. Further, Wu and Wang (2012) took things a bit further by trying to define a whole new set of gestures using hands and body movements. However,
as is so often the case, there is no real rigorous evaluative evidence provided for effectiveness and user experience. Despite these being rather early studies in this area, use of gestures is a method that would go on to be implemented by large corporations, e.g. by Samsung (2019). However, no rigorous or empirical data on usability evaluations of these kinds of televisions seems to be available.

In a different approach, Jovanov et al., (2015) used a 3D representation with animations for a television user interface. They presented various design concepts around a 3D theme. Although the authors suggest that what they have developed is good, no evaluation is presented to suggest that a 3D representation of the kind they are proposing would be better than what is on offer by current manufacturers.

Where some effort in addressing accessibility can be seen, is in the research of Oliveira et al. (2017). They looked at some existing products and developed a prototype. They then evaluated their prototype with nine participants and found that the participants on the whole were positive in relation to the prototype. However, we would suggest that it may have been interesting to have compared their prototype with one or more existing products.

This brief literature review indicates that much more rigorous effort needs to be done in evaluating any new prototypes or concepts providing more concrete evidence. More effort should also be carried out in evaluating existing commercial systems which are notoriously lacking in publicly available usability evaluations. The published literature further indicates that there is a lack of systematic and rigorous application of universal design principles in television user interfaces.

The next section aims to address one of the issues observed from the literature. This concerns evaluations of existing commercial products. Therefore, an evaluation using Nielsen’s Heuristics (Nielsen, 1994a) and the seven universal design principles (Story, 1998) is detailed for the Apple TV user interface.

### 3. Heuristic Evaluation

The current Apple TV user interface interacted with via the Siri remote control was firstly evaluated using the well known Heuristics by Nielsen (1994a). Broadly the evaluation followed the guidelines by Nielsen (1994b) regarding how to best conduct a Heuristic Evaluation. Three experts conducted the evaluation separately and then the findings were aggregated into one main result which is presented in the table below.

<table>
<thead>
<tr>
<th>Heuristics</th>
<th>Findings</th>
<th>Outcome</th>
</tr>
</thead>
</table>
| Visibility of System status | - On the homepage of the Apple TV, information was missing on the icons.  
- Only after hovering/navigating onto the icons would text be displayed describing the icon.  
- When selecting a certain icon, a new window was opened. However, there was nothing on the screen informing one which screen they were on, thus failing to display the system status. | Fail |
<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match between System and real world</td>
<td>- The icons used in the home screen and other screens somewhat depicted the characteristics of the real world. For example, for the Music functionality, the image of a half note of music was used.</td>
<td>Pass</td>
</tr>
</tbody>
</table>
| User Control and Freedom                     | - There is an absence of a back option in the user interface of all screens. To navigate back to the previous state, one had to navigate through a button and that button had multiple functionalities.  
- While navigating on the home screen, on the display screen, there was no sign or information for both horizontal and vertical scroll options.  
- While scrolling horizontally and vertically from the initial state, the only way to go back to the previous state was to scroll back.  
- While using the onscreen keyboard, to undo or correct, one had to go to the right most corner to select the backspace button. There was no alternative provided. | Fail   |
| Consistency and standards                    | - The icons were mostly presented in rectangular tiles as well as the content of the system. However, the sizes of the tiles differed screen by screen while navigating through the menus.  
- Some of the content/information were selectable and some were not. There was no distinct differentiation on the elements which were selectable and those which were not.  
- The tiles which contained the content were not of a consistent size in some screens. | Fail   |
| Error prevention                             | - While entering the email id for Apple id, there was no validity check for the correct email address.  
- If one accidentally navigated back to the home screen while entering the credential to register for Apple TV, then the form did not save the work in progress. | Fail   |
| Recognition rather than recall               | - Some elements were clickable and some were not, thus making it quite confusing to recognize if the similar elements had similar functionalities.                                                             | Partial Pass |
| Flexibility and efficiency of use            | - The interaction through the user interface was done via scrolling and clicking on the tiles of menus and the absence of shortcut menus create less flexibility in use.                                | Fail   |
A Preliminary Usability and Universal Design Evaluation of a Television App User Interface

- The Siri remote was the primary input device to use and had limited buttons and single button(s) have multiple functions.

<table>
<thead>
<tr>
<th>Aesthetic and minimalistic design</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The home screen had fewer menus aiding minimalistic design.</td>
</tr>
<tr>
<td>• However, while navigating through menus, some screens had hidden side menus, information on programmes with several elements on the screen.</td>
</tr>
<tr>
<td>Partial Pass</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Help users recognize, diagnose and recover from errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Due to focused based interaction, the user does not always realize how the menu will function before selecting them.</td>
</tr>
<tr>
<td>Partial Pass</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Help and documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>• There was no help section or manual on how to use the system or any particular screen.</td>
</tr>
<tr>
<td>Fail</td>
</tr>
</tbody>
</table>

### 4. Universal Design Principles Evaluation

Universal design essentially fosters a philosophy of design for everybody, ideally with no adaptation or specialized design ((Connell, Jones, Mace, Mueller, Mullick, Ostroff, Sanford, Steinfeld, Story and Vanderheiden, 1997) and (Story, 1998)). Universal Design advocates the accessibility and usability of a system for the users irrespective of their age and different abilities.

The Apple TV user interface was further evaluated using the principles of Universal Design ((Connell et al., 1997) and (Story, 1998)) by the same three usability experts that evaluated the Apple TV user interface using Heuristic Evaluation. Their findings after the evaluation are as follows:

**Table 2: Universal Design Principles Evaluation Results.**

<table>
<thead>
<tr>
<th>Universal Design Principles</th>
<th>Findings</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equitable Use</td>
<td>The apple TV consisted of features for people with physical disabilities, e.g. blindness and deafness. It had the feature of voice over for blind users and subtitles for deaf users. For viewers with colour blindness, they could adjust the colour contrasts and fields. Since the Siri remote was the primary device to navigate through the interface, the remote in itself was less accessible and usable in its term, so less usable for users with motor disabilities.</td>
<td>Has elements of equitability.</td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
<td>Result</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Flexibility in Use</td>
<td>Since the primary input device was a physical remote, it limits the ways of interaction for people with different abilities. The navigation was done through scrolling and clicking on the buttons, but the interface does not provide a clear state of the system.</td>
<td>Lacks flexibility in use.</td>
</tr>
<tr>
<td>Simple and Intuitive Use</td>
<td>The interaction of apple TV was based upon focused based interaction, which resulted in confusion for users as text information on an icon was shown only when that icon was selected. In some screens of apple TV, there were numerous elements where some were clickable and some were not. There was no distinct difference among such elements which were clickable or not. Users tended to seek help or more information about the technology they were using. In Apple TV there were no help features, the search option was displayed as the secondary function and was harder to access (one needs to scroll down or search for it).</td>
<td>Lacks in simplicity and intuitive use.</td>
</tr>
<tr>
<td>Perceptible Information</td>
<td>When a user navigated to a screen, there was no adequate information on which screen the user was currently viewing. Also, the sizes of similar elements of that screen varied from each other creating confusion for the user.</td>
<td>Lacks perceptible information.</td>
</tr>
<tr>
<td>Tolerance for Error</td>
<td>The forms in Apple TV did not have in-form validation checkers. Only after writing all the necessary information and confirming would show an error if any. There was an absence of in-form element error checking.</td>
<td>Lacks tolerance to errors.</td>
</tr>
<tr>
<td>Low Physical Effort</td>
<td>The on screen keyboard of Apple TV had an error correction (backspace) button at the right extreme. However, whenever one had to correct an error, one needed to traverse to the far right hand side to correct the error. There was no shortcut option for scrolling back to the original screen so one needed to scroll all the way to the top or to the left to get to the previous state.</td>
<td>Does not foster low physical effort.</td>
</tr>
<tr>
<td>Size and Space for Approach and Use</td>
<td>The Siri remote control can easily slip out of a user’s hands. At first glance it is not always easy to see if a user is holding the device with the correct orientation.</td>
<td>Lacks in size and space for approach and use.</td>
</tr>
</tbody>
</table>
5. Discussion of Results

As one can see from Tables 1 and 2, there are systematic failures at almost all levels with the user interface of Apple TV. Although not the main scope of this work, the Siri remote also indicates there to be design flaws in terms of usability and universal design.

The commonality with all the negative issues found in this two-part evaluation is that application of already pre-defined and existing user interface design guidelines have not been applied or have not been applied well enough. We feel that the issues are so basic in nature that they should have been ‘caught’ and designed out at earlier stages of the production.

The results indicate that there is a lack of understanding concerning usability for diverse users. Apple TV and similar products are likely to be used by people of vastly varying levels of ‘technical’ competence. Further, such products are likely to be used by a large spectrum of diverse ages.

The results also show that not much thought has been given regarding universal design which we would argue is strongly linked with usability and aspects of interaction with technology. A completely universally designed Apple TV would be ideally easily usable ‘out-of-the-box’ by anyone, irrespective of their abilities, age, gender and without the need for any adaptation etc.

5.1. Ways Forward

There are several suggestions we can make to improve television user interfaces. The first area of suggestion is that Apple does have its own set of guidelines (Apple Inc. 2020b) for television apps. Following these closely would somewhat help in resolving some of the issues discovered in the two-part evaluation described above.

However, some of their guidelines would need to be rethought. One example in the ‘App Architecture’ section under ‘Navigation’, Apple declare that a ‘Back’ button should not be displayed on the screen. They state that: ‘People know that pressing Menu takes them back, so don’t waste space in your app with an extra control that duplicates this behavior. (Apple Inc. 2020b)’ It therefore appears that Apple’s reasoning is about assumptions about users current knowledge and space saving on the screen. However, we would argue that this thinking is not in line with universal design principles. This statement makes assumptions that ALL users are using other systems that have a menu button that takes them back to some previous state. We would argue that this reasoning is flawed. The rationale given that having a back button wastes space is rather tenuous at best. Apps are potentially wasting on-screen space for other matters, so displaying a ‘Back’ button is hardly a problem. One particular universal design principle that is not fully met with this kind of design is the Flexibility in Use principle. There should be a choice of methods available where possible. Therefore, forcing users to ‘know’ that the menu button needs to be used is not flexible. Simple and Intuitive Use is another principle that is not met with this approach. It states that ‘Use of the design is easy to understand, regardless of the user’s experience, knowledge, language skills, or current concentration level (Story, 1998).’ Once again Apple’s approach here assumes that ‘everybody’ must know that the Menu button achieves the ‘Back’ action.

A further example concerns the ‘App Architecture’ section under ‘Authentication’. Apple begins this section by saying that ‘Apple TV is designed for entertainment, not data entry. Ask people to authenticate only in exchange for value,… (Apple Inc. 2020b)’.
Therefore, Apple is arguing this position from an ‘entertainment’ point of view. While this is true, we would argue that as soon as one asks the user to do any form of data entry, then the product also becomes a product requiring data entry. Therefore, the guidance given here to developers is not adequate enough (see above for issues concerning making corrections with data entry). Apple also advises developers to oblige (Apple uses the word ‘prefer’ here) users to do their authentication on a separate device. This presumably means something like a computer, smartphone or tablet etc. Although no clear reason for this requirement is given in this section of the Apple Guidelines, we would again argue here that this goes against universal design principles as it assumes users have other devices that are in easy reach to do the authentication. As soon as a user feel frustrated or feels that the authentication process becomes troublesome, the whole user experience suffers. The universal design principle of Equitable Use involves trying to avoid stigmatising users. Assuming users have one or more other devices for performing authentication could stigmatise some if they do not have other devices. Equitable Use also concerns making a design appealing. Being forced to pick up another device to authenticate by having a code sent by the television app is certainly not appealing. The third principle of Simple and intuitive Use also comes into play here. Forcing users to authenticate with some other device adds unnecessary complexity to the interaction.

The above two examples illustrate our argument that the developer guidelines would in several places require re-thinking and brought more in line with usability and universal design principles. More could be detailed, but for the sake of brevity we limit our examples to two.

The second area of suggestions for ways forward involves design using already well established usability and interaction principles. This option is often ignored or overlooked. For the sake of brevity we will illustrate this with one major set of principles. For several decades, Ben Shneiderman devised and refined over the years his ‘Eight Golden Rules of Interface Design’ (Shneiderman, Plaisant, Cohen, Jacobs and Elmqvist, 2018). These rules are a good basis to use and can also be modified to help with particular interface contexts.

Therefore with or without modification, we would argue these golden rules are applicable to Apple TV user interfaces. For example, Golden Rule 2 is ‘Seek Universal Usability’. This, amongst other things includes designing for the whole spectrum of users, from beginners to experts. In the Heuristic Evaluation detailed above, it was noticed in the unit we tested that there was no user manual or specific explanations available to a user. This clearly shows that Golden Rule 2 is not being followed in this aspect. Within the Apple design guideline sections we could not see any detailed guidelines on providing guidance to users. Under the ‘Onboarding’ section a very minimal amount of information is given to designers. Suggestion that a tutorial for beginners could be provided is included, but the suggestion indicates that good app design is better than a detailed tutorial (Apple Inc, 2020b). Apple’s idea seems to be that apps should be so intuitive that they do not need manuals. However, as mentioned, Apple has no information on how to use the actual streaming device etc. They possibly believe it is so simple that this is not required.

Another example concerns Golden Rule 5 – ‘Prevent Errors’. This is quite self-explanatory. However, it means that user interfaces should be designed as far as is possible to prevent all users from making serious errors. The interface should also be designed for easy recovery from an error. In the Heuristic Evaluation detailed above, it was found that there was no validity check for entering the correct email address in
relation to the Apple id. This clearly shows that Golden Rule 5 is not being followed properly. Furthermore, there was no clear information in Apple’s design guidelines concerning this issue. Perhaps Apple believes that good design will prevent errors anyway. However at the basic level, all advanced interface designers know this is not always the case.

The above two examples illustrate our argument that well established user interface design guidelines are being overlooked or ignored, thus causing the user experience to suffer.

The third area of suggestions for ways forward involves design with guidance from the International Organisation for Standardisation (ISO). ISO standards are usually developed slowly over a number of years (Bevan, 2006). Several ISO standards are available for user interface designers. These can be grouped into various categories. Bevan (2006) grouped them as follows: ‘1. The use of the product (effectiveness, efficiency, and satisfaction in a particular context of use), 2. The user interface and interaction, 3. The process used to develop the product, 4. The capability of an organisation to apply user-centred design’. Therefore, this indicates that there is a relatively rich set of ISO standards available for designers of user interfaces. However, many do not seem to take advantage of this rich set of standards when designing user interfaces (See the work of Murano (2018) for an example of how ISO standards could be used in user interface design).

Using the same example mentioned above, the Heuristic Evaluation detailed above, found that there was no validity check for entering the correct email address in relation to the Apple id. In ISO 9241-110: 2006 (International Organization for Standardization, 2006) there is a clear guide concerning validation and verification of data. It even states an example of an email client where it ought to verify that the syntax of an email address is correct. So this obviously indicates that this particular aspect has not been followed in the Apple TV user interface.

The final example we will mention concerns another issue found during the Heuristic Evaluation, where some of the content/information were selectable and some were not. There was no distinct differentiation on the elements which were selectable and those which were not. In ISO 9241-16:1999, section 5.2.2 covers ‘Distinctiveness of object representations and direct manipulation control icons’. This describes how designs should allow users to see which elements on the screen can be manipulated and which cannot be manipulated. The standard goes even further by saying that it should be clear to the user which types of direct manipulation can be used. This particular issue likely comes under the ‘App Architecture’ and ‘Focus and Selection’ section in the Apple design guidelines. This suggests to use ‘…subtle animations and the parallax effect (to) produce a feeling of depth that clearly identifies the item that’s currently in focus’ (Apple Inc, 2020b). However, this ‘model’ of interaction does not deal with the lack of differentiation in what is and is not selectable. Using focus interaction requires the user to explore everything (until memorised) by moving to each element and observing if the behaviour indicates something to be selectable or not. So this illustrates again how the application of an ISO standard could have improved this aspect of the interaction which was failed in our Heuristic Evaluation.
6. Conclusions

Overall, we suggest that proper and rigorous application of well defined standards and user interface design guidelines would strongly improve user interface designs and therefore are a good way forward. Most of these standards and guidelines have been around many years and so should be a natural and every day resource for all user interface designers and developers. As we have illustrated above, the faults discovered in our two-part evaluation would in most cases not have existed in Apple TV if well defined standards and user interface design guidelines had been used from inception.

The ways forward presented may appear to be simple in nature. However, we would suggest that often the user interface design community is ignoring or not fully putting these into practice.

We therefore suggest that this paper makes a very strong contribution to knowledge in two areas. The first, concerns that to our knowledge, no one has carried out such a detailed evaluation of the Apple TV user interface as documented in this paper. The results of this alone should provide the developers and designers with much needed information to improve their product. We would also suggest that to our knowledge very few detailed publicly available evaluations have been documented of any similar type products.

The second, concerns our illustrating how rigorous use of the ISO standards and Shneiderman’s ‘Eight Golden Rules of Interface Design’ (Shneiderman, et al, 2018) could have avoided what we believe to be basic design errors in the Apple TV user interface. This should alert designers to the usefulness of such standards and guidelines.

Since the indication is that media streamers such as Apple TV have not been given much attention in publicly available evaluations, possible future work should also evaluate other well known media streamers such as those on sale by Google, Amazon and others.

References


