Eye Tracking Evaluation of User Experience on Large-Scale Displays

Andrew Schall(✉)

Key Lime Interactive, Miami, FL, USA
andrew@keylimeinteractive.com

Abstract. Recent advancements in technology have made eye tracking less expensive, much easier to use, and flexible enough to track a variety of display sizes and configurations. Larger high-resolution displays have become an increasingly prominent format for many users. New user behavior patterns have been emerging between primary and secondary (also known as second screen) displays. This paper describes a new research approach in order to understand what attracts user attention and identifies what they see when interacting with these devices. A case study is presented that demonstrates the procedures and findings for a study that involves eye tracking of a large-screen television display. The study described is a user experience evaluation of dynamic on-screen content presented as a part of the display during a television program.

Keywords: Eye tracking · Large-scale displays · Television · User experience · Second screen · Study design · Case study

1 Introduction

Eye tracking is now more accessible to UX researchers than ever before. Recent advancements in technology have made eye tracking less expensive, much easier to use, and flexible enough to track a variety of devices. However, few UX researchers are aware that eye tracking isn't just for computer screens anymore. Eye tracking technology and its uses are evolving. The ability to accurately and unobtrusively conduct eye tracking research on large-scale displays was nearly impossible until only very recently. Larger high-resolution displays have become a dominant interface for our users and we need to be able understand what attracts their attention and what they see when interacting with these devices.

Conducting eye tracking on large displays can be complex. Eye tracking needs to be carefully considered during the planning of a user research study. Testing these displays with eye tracking can be daunting if the study objectives are not carefully selected, and if plans have not been made to make the necessary accommodations required to obtain reliable and accurate data.

Large display experiences are highly contextual and it can be difficult to recreate the user's environment within a lab setting. It is critical to establish the optimal distance between the user, the eye tracker and the display being evaluated. This can be complex

M. Antona and C. Stephanidis (Eds.): UAHCI 2015, Part I, LNCS 9175, pp. 98–108, 2015.
DOI: 10.1007/978-3-319-20678-3_10

due to limitations of the eye tracking hardware which often requires a relatively short distance between the user and eye tracker. This distance can also impact the trackability of the display itself. The larger the screen, the greater the needed distance between the participant and the display.

New user behavior patterns have been emerging between primary and secondary (also known as second screen) displays. These second screens often display content unrelated to what is being shown on the larger display and can be a source of distraction for the user. Researchers need to understand the tradeoffs associated with allowing participants to use a second display, which tends to be more natural, versus the ability to collect as much eye tracking data from the primary display as possible. It is important to understand the different eye tracking configurations available to create a balance between a realistic environment and the need to collect comparable data across participants.

Other considerations for conducting eye tracking studies with large displays include carefully planning out areas of interest for analysis. Manual segmenting of video clips can be necessary when comparing highly dynamic media across different screen regions across many participants. Large high-resolution displays also require additional computing power for capturing, analyzing and storing eye tracking data.

A case study is presented that demonstrates the study design, lab configuration, and analysis procedures for a recent study involving eye tracking of a large-scale display. The study involved the user experience evaluation of dynamic content presented as a part of the display during a television segment.

2 Using Eye Tracking to Measure User Experience

In nearly all cases, the user experience of digital interfaces is driven by visual output. These days, user experience designers are creating visual content for everything from wall-mounted displays to laptops to mobile devices.

Our visual field is constantly being bombarded by many concurrent stimuli. We are overloaded and overwhelmed by visual information, and we constantly resort to prioritizing what we pay attention to. To measure the effectiveness of content, researchers need to determine what users are looking at and what they choose or do not choose to engage with [1].

The user experience of television has been studied for numerous decades and more recently through the use of eye tracking. A fundamental difference between television and other digital media is that it tends to be more passively experienced. This puts greater weight on the viewing rather than interacting experience. User engagement becomes less about what the user is doing and instead becomes predominately about what they are visually engaging with. This makes eye tracking a natural fit for studying the user behavior of television programs.

2.1 Related Work

Eye tracking has been used to evaluate the user experience of websites to understand how users perceive and work the interface. However, eye tracking has been used in little research on television [2]. Several studies have explored the presentation of on-screen information overlaying television programs including research by Josephson and Holmes in 2006.

Researchers have found3 that individual looks at the TV vary in length and people develop different watching strategies to follow content on TV. For example, people may look at the TV only at the right times, just enough to be aware of what is happening, while being engaged in some other activity.

In the late nineties Jakob Nielsen published an article [4] based on his work comparing the experience of watching television with interacting with a computer. The diagram below summarizes the key differences between these two mediums.

	Television	Computer
Screen resolution (amount of information displayed)	relatively poor	varies from medium-sized screens to potentially very large screens
Input devices	remote control and optional wireless keyboard that are best for small amounts of input and user actions	mouse and keyboard sitting on desk in fixed positions leading to fast homing time for hands
Viewing distance	several meters	a few inches
User posture	relaxed, reclined	upright, straight
Number of users	social: many people can see screen (often, several people will be in the room when the TV is on)	solitary: few people can see the screen (user is usually alone while computing)
User engagement	passive: the viewer receives whatever the network executives decide to put on	active: user issues commands and the computer obeys

Nielsen highlights key differences between the two devices that imply how the user is likely to experience and interact with each medium. While this article was written prior to the widespread adoption of smartphones and tablets it helps to establish basic operating parameters that guide how television studies should be conducted.

Brown et al. [5] take this a step further by applying these facts to the design of a user research study. TV viewing typically occurs in a relaxed environment, quite different to that of a typical usability lab. This environment complicates experimental setup compared to an office/desktop computer scenario, with even basic challenges such as viewing distance potentially making data capture difficult.

2.2 Understanding Eye Movement Behavior While Watching TV

Holmes et al. [6] compiled key findings from several researchers who studied visual attention of the television experience using eye tracking. Based on this research, it was determined that the amount of uninterrupted sustained eye gaze on the television is only about 7 s long at a time. An even shorter period of time (less than 2 s) reflects active, informed monitoring of content that they equate to “checking in” for those familiar with the program.

Financial and news programming contains an assortment of ever changing content.

This type of viewing for frequent, yet very short periods of time also makes sense for those watching financial and news programming where the information changes very quickly. Viewers of this particular type of content are used to obtaining small fragments of information at a time with frequent periods of looking towards and away from the screen.

2.3 Use of a Second Screen

Many users today do not sit down and entirely focus their attention on the television screen. The ubiquitous nature of mobile devices has created the phenomena of a two-screen experience. In 2014, 84 % of smartphone and tablet owners said that they use their devices as second-screens while watching TV at the same time [7]. To better understand the second-screen experience Holmes et al. performed an eye tracking study where participants were asked to watch a program on a television and also use a companion app on a tablet computer. They found that on average only 63 % of the participants’ attention went to the television during the program, and 30 % to the tablet (and 7 % off of both screens).

The effects of a second device should be a major consideration for any studies involving the study of real world television viewing experiences.

3 Case Study: Evaluating the Effectiveness of a Financial TV Segment

Many news and financial television networks utilize on-screen visualizations complementary to their standard programming. These can include news alerts, stock performance, and additional details related to a story. One particular organization, which shall be referred to as "Financial Network 1" for the purposes of this paper, wanted to better understand the viewing behaviors of their audience. The executives and producers of the network's programming had several assumptions related to typical viewing patterns and areas of attention. They believed that by providing a screen with less visually complex information it would encourage viewers to engage more with the content displayed. They also believed that the type of content displayed is the information that their audience most wanted to see.

The research team involved also theorized that many of the audience members would likely multitask while watching the network, which may include the use of smartphones, tablets, laptops and non-digital media as well. The team wanted to be able to understand typical viewer behavior while they watched live television as opposed to prerecorded programs.

3.1 Research Goals

A series of research objectives were established in order to better understand user behavior while watching the network. The goals included:

- How quickly do they notice each of the on-screen elements?
- How long do they spend looking at certain Areas of Interest (AOIs)?
- How many times do they look at an AOI during the viewing period?
- Do they read the bullets and headlines? How many do they read?

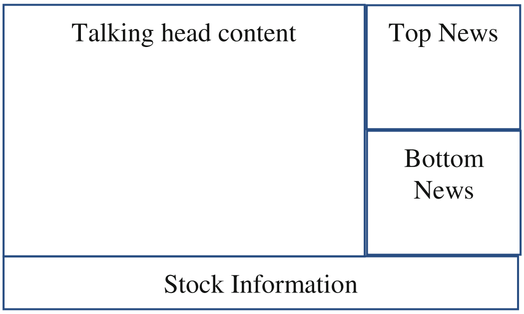
Another goal of the study was to compare the results with a competing television network's programming to see if the different design layout contributed to the consumption of different types of information.

3.2 Stimulus Materials

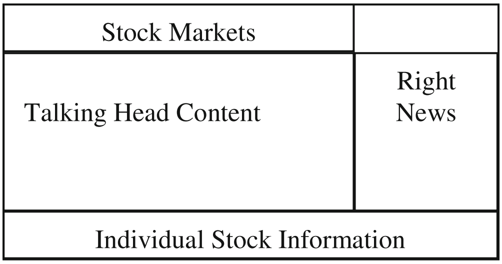
All of the network programming contains the same types of onscreen elements such as a dedicated box for news stories and stock information. The various elements of the on-screen displays were categorized into a series of Areas of Interest (AOIs) that were then later used during analysis.

3.3 Participants

Given the distinctive type of content provided on both television networks it was critical to obtain participants that would normally watch financial news programs. This included both members of the general public with personal portfolios as well as professional investors who manage the portfolios of their clients. Participants were asked about their current viewing behaviors including which specific networks they watch and how often they watch them. Our study included a total of 35 participants.



Layout of Financial Network 1.



Layout of Financial Network 2.

3.4 Test Protocol

All participants watched live TV programming, so they saw different programming depending on their session times. All sessions took place between 7 am and 6 pm. Viewers were asked to simulate their normal TV viewing behavior by using the desktop PC or their own personal devices (e.g. phones, tablets, laptops); they were not required to watch TV the entire time.

Participants watched each channel, Network 1 and Network 2, for 15 min (total of 30 min TV viewing). The order of the two channels was alternated (e.g., P1 watched Network 1 then Network 2, P2 watched Network 2 then Network 1, etc.) in order to eliminate any order bias. Each session in total lasted approximately 45 min to 1 h

After watching both networks, participants were asked a series of questions about their viewing experience. This provided a qualitative perspective on why participants were interested in certain content and areas of the display.

The research team made efforts to create as natural an environment as possible for participants. The setting used was meant to emulate a typical home office or desk at an office. The environment included commonly found elements such as a desktop computer, telephone, large working surface, and also a television.

Time (1 hour in length)	Section Description
0:00-0:10 (10 minutes)	Section I. Introduction and background questions • What do you do for a living? • When are you normally watching TV? • What do you do while watching TV? • Why do you watch and/or
0:10-0:15 (5 minutes)	Section II. Eye Tracking Calibration
0:15-0:30 (15 minutes)	Section III. Live video segment #1 • / or (randomized) • Allowed to multi-task and use computer and/or personal device
0:30-0:45 (15 minutes)	Section IV. Live video segment #2 • / or (randomized) • Allowed to multi-task and use computer and/or personal device

0:45-1:00
(15 minutes)
	Section V. Qualitative feedback and follow-up questions
	• Overall feedback regarding on-screen content and news programming
	• Ranking screen content areas
	• Comparative rankings between and
	  ◦ Stock and market information
	  ◦ On-screen readability
	  ◦ Overall usefulness of screen data

Test session breakdown.

Participants were encouraged to use their own personal devices throughout the test session including tablets (e.g. iPads) and smartphones (e.g. iPhones). They also were given access to a desktop computer with dual monitors in order to browse the Internet.

1. Participants viewed media on a 46-inch TV.
2. The eye tracker was placed on the desk in front of viewers.
3. A computer was available for participants to use while watching TV.
4. A ceiling mounted camera was used to capture the participant's face

One of the most challenging technical aspects of the study setup involved the eye tracker itself. We used a Tobii X2 Eye Tracker [8] to track the participant's eyes while looking at the television monitor. The eye tracker can only do accurate tracking of a person's eyes if the visual angle, as seen from the person's eyes, does not exceed 36°, between the center of the eye tracker and any point on what the person is looking at. This is valid as long as the person stays within the area in front of the eye tracker defined as the area of freedom of head movement, which is roughly between 40 and 90 cm from the center of the eye tracker when seen from the side. This means that depending on the position of the eye tracker in relation to the display, it can track different sizes of displays. If the display is positioned further away from the user, a larger area can be tracked.

Another feature is the aspect ratio of the display. Since the optimally tracked area is about half circle, the eye tracker can track displays with a large width to height ratio better than displays where this ratio is small. This was ideal for our setup, which included a typical widescreen HD television display.

The limitations of the technology required that we place the eye tracker on a small tripod directly in front of the participant with the television display set up approximately 1.8 m further away. The setup was also limited by the distance the television could be from the participant. In an ideal setting, the television would have been wall-mounted and placed a greater distance from the participant, however this was not possible due to the limitations of the lab space.

3.5 Analysis

The analysis included a qualitative aspect focusing on general user viewing patterns and areas of engagement as well as a quantitative aspect focusing on predefined areas of interest.

Due to the nature of eye tracking user interfaces over video content, traditional automated and aggregated eye tracking analysis methods (such as heat maps) could not be generated for the television interfaces tested. Instead, we relied on a more time-consuming qualitative analysis method: watching videos of each session. While watching each session, we looked for recurring patterns of usage behavior, paying special attention to participants' eye fixations and saccades.

3.6 Results

Not surprisingly, the most amount of viewing time was spent watching the primary news story (72 %). However, a significant amount of time was spent viewing the other onscreen information. It was determined that over a quarter of viewing time (28 %) on

Network 1's programming was spent reading news content on the side of the screen. This area was separated out into two areas. The top section received 20 % of attention while the bottom half received only 8 % of viewer's attention. Participants also looked at the top news section more frequently (every 13 s compared with the bottom (every 27 s). By using the time to first fixation metric we were able to determine that participants noticed the news area early on in their viewing experience. On average, they first viewed the top news area within the first 13 s and the bottom news area within 18 s. In comparison, it took participants much longer to notice other onscreen elements such as the stock ticker (42 s), market data (46 s) or date/time (60 s).

Measure	Definition	What it indicates?	Unit of measurement
Total Visit Duration	Cumulative amount of time spent looking at a particular area	Fixed attention and attraction to a given area	seconds % of total time spent
Visit Count	The number of times a person looks at a particular area	Usefulness and attraction to a given area	# of times visited
Visit Duration	The average length of time per look in a particular area	Fixed attention and usefulness of a given area	seconds
Time to First Fixation*	The amount of time before a person looks at a particular area	Noticeability of a given area	seconds

**Fixation is a pause in eye movement. Saccade is rapid eye movement between fixations.*

On average participants looked at the main story most frequently (11 s). In comparison, the news headlines attracted their attention approximately every 19 s. Participants looked at the market data the least (53 s).

In our post experience debriefs participants said that they found the news on the right side to be informative, in proportion to their interests, and that the information stayed on the screen long enough to understand what the story was about. The real-time eye gaze data supported this by showing clearly defined left to right reading patterns across each of the bullets of information displayed in the news box. Most participants appeared to read at least two of the bullets before the content changed. We were surprised by how quickly participants were able to scan over the content. The average viewing time for the top news box was only 2 s on average per visit, however this was sufficient time to read the contents of a bullet or a headline.

Participants were least interested in viewing the market data. During the debriefs participants said that they only wanted to see general market trends such as whether the major markets were trending up or down. Most were not interested in seeing individual stocks because they would normally look this information up on their computer and would not want to wait for a specific stock to appear on the TV screen. This was supported by the eye tracking data that showed attention on the market data to last no more than a second or two at a time.

The news information performed better overall on Network 1 than on Network 2. Participants spent 61 s longer reading news content on Network 1 than Network 2 (a total viewing time of 28 % compared with 15 %). However, out of the total viewing time participants spent more time looking at Network 2's stock information (46 s compared to 25 s total).

4 Conclusion

Recent advancements in technology have made eye tracking less expensive, much easier to use, and flexible enough to track a variety of devices. Eye tracking technology and its uses are evolving. Larger high-resolution displays have become a dominant interface for our users and we need to be able understand what attracts their attention and what they see when interacting with these devices. Eye tracking needs to be carefully considered during the planning of a user research study. Large display experiences are highly contextual and it can be difficult to recreate the user's environment within a lab setting. New user behavior patterns have been emerging between primary and secondary displays. Researchers need to understand the tradeoffs associated with allowing participants to use a second display, which tends to be more natural, versus the ability to collect as much eye tracking data from the primary display as possible. It is important to understand the different eye tracking configurations available to create a balance between a realistic environment and the need to collect comparable data across participants.

References

1. Bergstrom, J.R., Schall, A. (eds.): Eye Tracking in User Experience Design. Elsevier, Waltham (2014)
2. Ali-Hasan, N.F., Harrington, E.J., Richman, J.B.: Best practices for eye tracking of television and video user experiences. In: Proceedings of the 1st International Conference on Designing Interactive User Experiences for TV and Video, pp. 5–8. ACM, October 2008
3. Vatavu, R.D., Mancas, M.: Visual attention measures for multi-screen TV. In: Proceedings of the 2014 ACM International Conference on Interactive Experiences for TV and Online Video, pp. 111–118. ACM, June 2014
4. Nielsen, J.: TV meets the web. Alertbox for February 15 (1997)
5. Brown, A., Evans, M., Jay, C., Glancy, M., Jones, R., Harper, S.: HCI over multiple screens. In: CHI 2014 Extended Abstracts on Human Factors in Computing Systems, pp. 665–674. ACM, April 2014

6. Holmes, M.E., Josephson, S., Carney, R.E.: Visual attention to television programs with a second-screen application. In: Proceedings of the Symposium on Eye Tracking Research and Applications, pp. 397–400. ACM, New York (2012)
7. What's empowering the new digital consumer? Nielsen.com, http://www.nielsen.com/us/en/insights/news/2014/whats-empowering-the-new-digital-consumer.html. Accessed 1 Feb 2015
8. Tobii X2-60 Eye Tracker User's Manual V.1.0.3. Tobii, 2014. Web, 1 February 2015