



WD201 Web Documentary

Eye-tracking Report

A user study on a Multimedia
Storytelling Experience

CONTENT

01

Introduction

02

Research Design

Stimulus Material

Participants

Research Goals

Methods

03

Main Findings

Usability

Pedagogical Functionalities

Visual Attention

04

Suggested Improvements

05

Conclusion

01 INTRODUCTION

Content – our new reality. We surround ourselves with it; from news articles to social media posts, from documentaries to VR experiences. But how do users interact with the constant stream of information they are subjected to? Which elements consume the most of their attention? How do they engage with them?

These questions are fundamental to comprehend the true potential of more immersive forms of media – especially, when taking “User Experience” into consideration. Can their experience be evaluated by using one of their own senses?

Eye-tracking technology has been a useful tool for conducting research on human attention for decades.¹ The same processes can determine User Engagement and perception of VR experiences and interactive documentaries. Particularly, due to the decreasing attention span of humans in times of media overload, and due to new user habits in general, there is a huge potential to use eye-tracking as a method to focus more on the individual reception of users.² By analysing that data, media professionals can obtain findings to improve formats such as web documentaries and VR experiences.

Regarding our own eye-tracking research, we took a closer look at the Project 6x9: *A virtual experience of solitary confinement* – Guardian’s first virtual reality experience. Our eye-tracking report begins with the research design, more precisely the presentation and justification of the stimulus material, the participants, the specific research goals and the used methods. After presenting our main findings with regard to usability, pedagogical functionalities and visual attention; we completed the report with our suggested improvements and a conclusion.

¹ Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. OUP Oxford.

² Vázquez-Herrero, J. (2021). Enhanced Experiences in Interactive Nonfiction: An Experimental Study on the Effects of Nonlinearity and Interactivity. *International Journal of Communication* 15: 1414-36
<https://ijoc.org/index.php/ijoc/article/view/16145/3388>

02 RESEARCH DESIGN

Stimulus Material

For our study, we used the VR experience 6x9 of the Guardian. It is a piece of interactive documentary reflecting on the incarceration and solitary confinements in American prisons. The main reason to use such a format is that we are working on our own VR project and thus we can incorporate the findings of our study directly into our work. Additionally, the Guardian is a leading player in interactive storytelling projects. Unfortunately, we realized that we do not have the technical equipment to record the user's gaze through VR glasses. Nevertheless, we used the 360° video of the project to conduct our research – with a computer as the output source.

Participants

A total of four people took part in our study. It was important to us that the participants were as diverse as possible. For example, other studies have shown that older people in particular have a different user behaviour when it comes to new technologies.³ For this reason, we wanted to include at least one person from the older generation. At the same time, especially people from younger generations are using web documentaries.⁴ The research with the participants was conducted on 11th of October 2023 using Tobii Pro Glasses 3 system.

Participant 1

female
in her early 20s



Participant 2

female
in her mid 20s



Participant 3

male
in his late 50s



Participant 4

female
in her mid 20s



³ Oh, Y. H., & Ju, D. Y. (2020). Age-Related Differences in Fixation Pattern on a Companion Robot. *Sensors*, 20(13), 3807. <https://doi.org/10.3390/s20133807>

⁴ Ducasse, J., Kljun, M., & Pucihar, K. Č., (2020). Interactive Web Documentaries: A Case Study of Audience Reception and User Engagement, *International Journal of Human-Computer Interaction*, 36 (16), 1558-1584, <https://doi.org/10.1080/10447318.2020.1757255>

Research Goals

Before we decided what methods to use, we clarified what we actually want to achieve with our study. Therefore, we formulated the main questions the research will confront and strives to answer:

- Where do the users look first?
- Which elements get the most attention? What do they look at the most?
- How do they orientate themselves when put in a new virtual environment?
- Do the users recognize all items and elements? If not, which ones were not recognized and why?
- How do the voices and ambient noises influence the experience?
- How much attention do they spend to the text and to the objects?

Our main goal is not to conduct research on the content of the experience – rather, we want to look at the specific user behaviour, in particular the orientation processes and behaviour patterns. In this way, we can then gain important insights to build our own user-friendly virtual world.

Methods

For our study, we first decided to use the think-aloud protocol. This method is one of the most widely employed usability methods and is particularly suitable for researching specific user behaviours⁵ – which fits perfectly with our research goals. We told the participants that they could comment on whatever they want during the experience. We hoped to track the feelings and reactions of the users in real time.

We tested this method on ourselves first and decided to not give any information about the VR project to the participants before. However, what we did not realise is that this method does not offer us any possibilities to ask about exciting observations. Therefore, after the first participant, we added another method to our research design. We also conducted qualitative interviews with the participants afterwards to discuss what they have described during the experience. With those individual questions on top we were able to gain even more profound insights. Although it is not ideal to adapt the methods after the launch, this is not uncommon in scientific practice and even required if it improves the quality of the research.⁶

⁵ Boren, T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. IEEE transactions on professional communication, 43(3), 261-278.

⁶ Hesse-Biber, S. N. (2010). Mixed methods research: Merging theory with practice. The Guilford Press.

03 MAIN FINDINGS

We analyzed the participants' gazes through the eye-tracking glasses, incorporating the findings of the think-aloud protocol and the qualitative interviews. Overall, the results can be classified into three categories.

Usability

Our first important finding is that almost all participants had problems with navigation at the beginning. Either, because they did not understand how to handle the control (especially the zoom in), or because they did not realise that they are an active part of the experience. One participant needed more than 20 seconds to notice that he was able to navigate. Looking at the heat map of the usage instruction (Figure 1), it can be seen that not a single participant looked at the navigation icon in the upper left corner. The poor presentation of this button and the fact that there was no further information about it are the main reasons for this lack of usability. Even when the experience itself begins, the focus is only on the content and not on the button explaining the navigation (Figure 2).

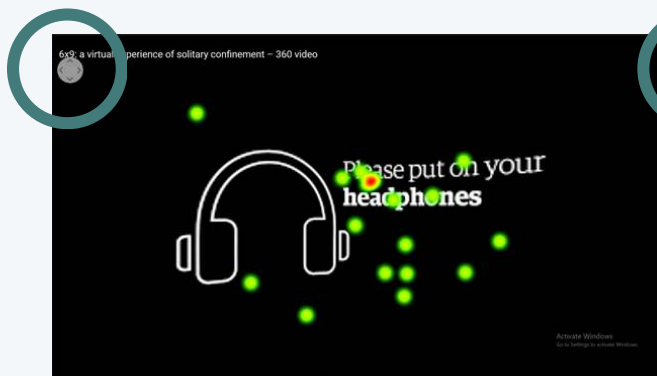


Figure 1: Heat Map of the usage instruction

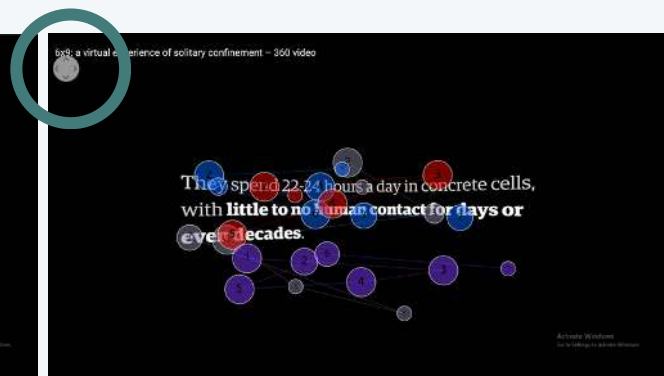


Figure 2: Gaze Plots of the content introduction

Pedagogical Functionalities

Our focus in exploring the pedagogical functionalities was that we wanted to find out when the content had a negative impact on usability. Fundamentally, the content was understood by all participants. Problems with the communication of the content occurred when too many stimuli impacted on the users at the same time.

The combination of text on the wall, narration in the background and ambient sounds was considered confusing by three out of four participants. This overload of stimuli had a negative effect on the comprehensibility of the content. At the same time, the basic idea of conveying the content in many different ways (text, narration...) was rated as very positive – as long as it did not take place simultaneously.

It was also said that the amount of text during the experience was sometimes too much. It would be better to provide more context before the start, so that only the absolute minimum of text needs to be used during the course of the exploration. One participant found it confusing that imaginary text was projected on the wall in a real-life environment.

Visual Attention

The most important question for us was how the participants orient themselves in a new environment. The analysis of the gaze plots in the area of interest (Figure 3) show that two elements in particular are looked at immediately. On one hand, these are the objects that are in the experience (in this case the toilet), and on the other hand, the light sources. The light gap in the door is perceived by all participants with at least the fifth gaze, which is a very interesting finding. At the same time, it can be seen that the participants' gazes move quickly from left to right when they are unsettled, mainly scanning the central area.



Figure 3: Gaze Plots of the moment the participant was put into the VR world

Furthermore, we have determined in our research design that, regarding visual attention, we want to find out whether, and for how long, different elements are perceived. Therefore, we will take a closer look now at text, objects and sounds.

The analysis of the eye-tracking data shows that the texts on the wall are almost always perceived when there is nothing else in the environment (Figure 4). The texts are looked at intensively and completely. All text snippets are scanned one after the other. Most of the time, the users start with the small snippets of text and if they have enough time before another stimulus comes along, they also read through the larger amounts of text. It becomes exciting when text and objects can be seen simultaneously. When analyzing these gaze plots (Figure 5), it is clearly visible that the objects get more attention than the texts. For example, all eleven gazes of participant 3 are scanning the objects, not a single gaze falls upon the texts.

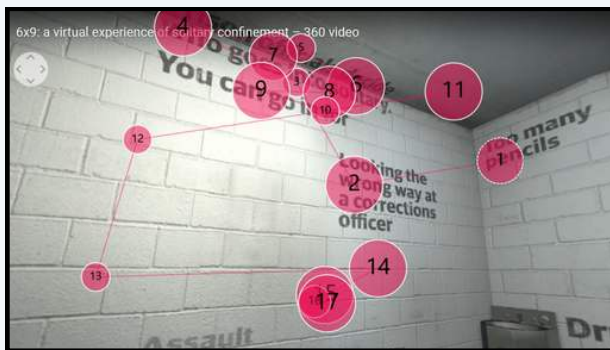


Figure 4: Gaze Plots of the texts on the wall

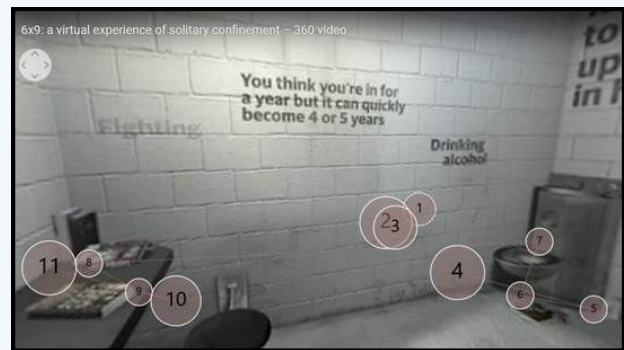


Figure 5: Gaze Plots of the objects in the room

Furthermore, the analysis of the cover image (Figure 6) shows that text highlighted in colour, and especially symbols integrated into the text get a lot of attention. As with the usage introduction, it can be seen that the elements in the upper corners do not get any attention – even if the user looks at it for more than ten seconds.



Figure 6: Heat Map of the cover image of all participants

The ambient sounds were rated as very important and impressive by all participants. The sounds helped the participants feel like they were actually in the room. The fact that the sounds came from different directions made the experience very realistic.

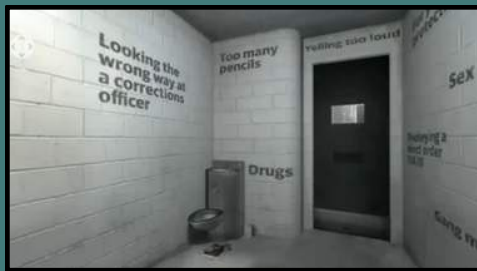


In the course of the study, we learned an important lesson. Since we used a 360° video as a stimulus, it is up to the participant to decide which parts of the experience they want to look at. However, in order to create heatmaps with several participants, the exact same snapshot is needed. Our users explored the 6x9 project in very different ways. For that reason it was not possible for us to create heatmaps with more than two users – apart from the introduction, which is the same for all of them. This limitation is an important lesson we learned.

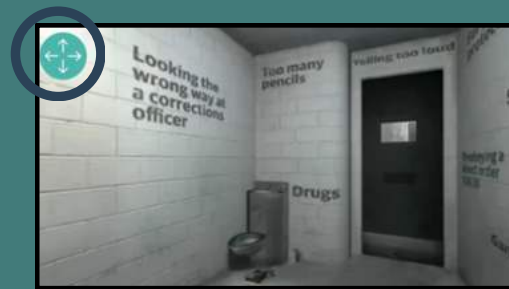
04 SUGGESTED IMPROVEMENTS

Based on feedback and eye movement analysis, suggested improvements are:

- Exploring a 360° video can be difficult for a user who has never experienced this type of format. One of our findings was that the users did not understand when and/or how to use the control.
 - First of all, to solve this problem, the navigation button displayed on the video screen could be more prominent, more visually attractive, to allow the user to understand how to use the 360° view.



Initial navigation button of the project

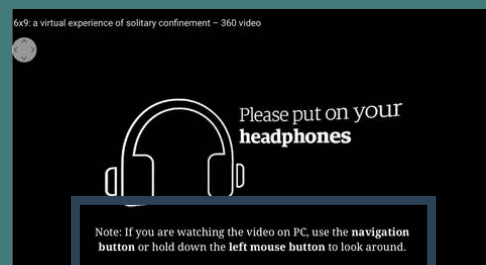


Possible improvement for the navigation button

- Second of all, another solution could be to add, at the beginning of the video, an explanation, through icons or highlighted text, of how users should navigate the project.



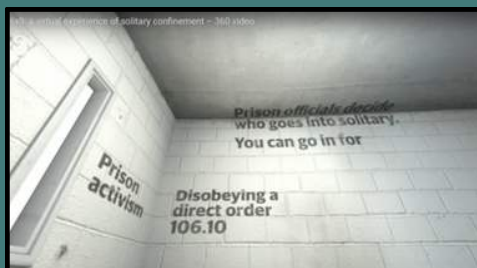
Initial usability introduction of the project



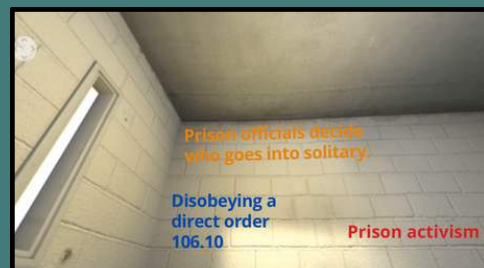
Possible improvement for the usability introduction

- Another improvement regarding an easier orientation is the lighting. It could be seen that light sources attract a lot of interest. A promising possibility is to use the lighting as a means of orientation. For example, one could illuminate the parts of the room that are to be seen first differently or brighter.

- To face the overload of simultaneous stimuli used in the 360° video, one solution could be to synchronize the narration and the readable text. With this change, it becomes easier for the human brain to focus on the experience and the user can perceive the content in a better way. Another improvement could be to place more context before the start of the experience, so that there is less text in the environment and other effects like the narration or ambient sounds have a greater effect.
- As a last improvement, the texts written on the walls have to be recognized when displayed at the same time as objects. This could be achieved in many ways – for instance by using (complementary) colours and contrasts or by throwing more light on the text.



Initial text snippets on the walls



Possible improvement for the texts on the walls

05 CONCLUSION

After successfully conducting our research of the 6x9 project of the Guardian using different methods, the following conclusions can be drawn regarding the development of future Multimedia Storytelling Experiences:

- An adequate **explanation of the interface** is the basis of every experience. If done successfully, can the user concentrate fully on the content and the visual elements.
- When users enter new virtual environments and want to orient themselves, they focus primarily on **objects** and **light sources**.
- Objects have a higher visual appeal than text. They catch the user's eye first and are looked at for a longer period of time.
- To make the text more attractive to users, a combination of **highlights, colours** and **built-in icons** can prove useful. It can also help to nudge the user by using different lightings in the room. Small text snippets are looked at first, larger amounts of text are only seen afterwards.
- Ambient sounds and narrations are seen as **very positive and helpful** by the users – as long as those auditory stimuli do not clash with the visual ones and lead to an overload of information.

The majority of our questions determined in our research design could be answered with our eye-tracking research. As a next step, it is important for us to take these findings into account when developing our own interactive experience.



HØGSKULEN I VOLDA