KEY CHARACTERISTICS OF VR VIDEOS AND THEIR IMPACTS ON AUDIENCE ENGAGEMENT

by

MADELINE WILLSON

Broadcast Journalism Major, Minor in Graphic Design

A FINAL HONORS PROJECT

submitted in accordance with the University Honors Program requirements

A.Q. Miller School of Media and Communication
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

May 2024

Approved by:

Huyen Nguyen

Abstract

This content analysis of 300 YouTube VR videos took a quantitative approach to examine the impacts of genre, interactivity, storytelling richness, technical enhancement, technical disruption, and immersive audio on audience engagement. Previously, there have been few quantitative studies that have focused on more than one or two of these categories in relation to 360-degree videos. This study is one of the first to consider the impact of multiple categories on audience engagement solely for YouTube VR videos.

Viewers are more engaged with VR videos that did not have technical disruption, as well as VR videos that had a higher level of storytelling richness. The impacts of genre, interactivity, technical enhancement, and immersive audio had no significant impact on audience engagement. This is consistent with previous studies as YouTube is significantly limited in regards to VR related features. Findings from the study contributed to the discussion on the technical behavioral aspect of human-to-content interactions, a part of the newly proposed audience engagement theory currently nested under the broader Uses and Gratifications framework. Future studies should consider using videos across the YouTube platform, as our study focused only on videos found on the YouTube VR channel.

Keywords: VR video, audience engagement, interactivity, technical disruption, immersion

Key Characteristics of VR Videos and their Impacts on Audience Engagement

I. Introduction

This study was inspired by Ekström's (2002) three-pronged framework to discern the epistemologies of a medium. The framework suggests researchers study the form of knowledge, production of knowledge, and public acceptance of knowledge claims. Based on this framework, we delve into the key characteristics of 360-degree virtual reality (VR) videos and how viewers engage with this specific format of immersive experiences. VR videos are known for their spherical viewing experience, where users can navigate around a virtual space in any direction. Key characteristics focused on in this research are genre, immersion, interactivity, and storytelling richness, which were previously highlighted in immersive storytelling studies (Wang, Gu, & Suh, 2018; Elmezeny, Edenhofer, & Wimmer, 2018; Uskali, Gynnild, Jones, & Sirkkunen, 2021; Pavlik, 2021; Zhang, 2022). Audience engagement is measured by viewer statistics displayed along with each video, which is collected through the use of YouTube Data API. This was suggested in a prior YouTube video study (Xie, Damiano, & Jong, 2022). In the media economy, audience metrics are "currency" (Steensen, Conill, & Peters, 2020). Emerging media formats and their producers need this type of "currency" to survive and create more compelling content. Therefore, our quantitative study has the potential to suggest areas for VR video production improvement. This approach has not been well-studied by media scholars, who more often conducted qualitative or experimental studies. The present study is also one of the first to delve into various media genres instead of just focusing on music or journalism. Additionally, our study aims to contribute to the line of uses and gratifications (U&G) studies which argued

that viewers might "unconsciously" (Khan, 2017) manifest their true motives and gratifications sought via audience metrics.

In brief, we first conducted a content analysis of 300 randomly selected videos from a population of 1,323 VR videos published on the official YouTube VR channel between 2015 and 2022. Then, we examined the relationship between each key characteristic and the level of audience engagement, controlling for the time of publication. As YouTube is a mature platform for immersive videos with a large user base, its VR videos are very suitable for quantitatively studying characteristics and technical-behavioral audience engagement. Potential societal and economic benefits, as well as emerging issues that hinder the power of transforming human behaviors and sharing cultures of immersive technologies, will also be discussed at the end of this paper.

II. Literature Review

A. Characteristics of immersive videos

a) Genres: Discussion on VR video genres and genre classifications has been very limited and often inconsistent (Wang et al., 2018; Foxman, Beyea, Leith, Ratan, Chen, & Klebig, 2022). The classification of VR video genres can be based solely on content (Bleumers, Van den Broeck, Lievens, & Pierson, 2012), real and spatiotemporal factors (Uksali et al., 2021), or a combination of content and platform (Foxman et al., 2022). Focusing on journalism, Uskali et al. (2021) classified 360-degree VR videos into four categories – live, news, documentary, and fiction – based on real time/real space or constructed time/constructed space. Assuming common features between VR videos and video games, Foxman et al. (2022) developed a complicated classification system based on traditional gaming genres, content, and hardware

support platforms. Both methods are not very practical when it comes to classifying VR videos published on YouTube. In reality, YouTube has put VR videos into various playlists that go beyond the categories listed by academic scholars. There are other categories such as gaming, tour, eyewitness, thrillseekers, music, art performance, vlog, winter sports, or just simply tranquil videos. Indeed, it's almost impossible to classify VR videos based solely on the subject matter because this emerging medium integrates diverse features of other mediums and does not seem to have a restriction on any subject matter. For practical purposes, our study joins the genre discussion by classifying YouTube VR videos into two categories – fiction and non-fiction, with content-based subcategories for both genres. This classification was primarily based on the realism factor adopted from Uksali et al.'s (2021) study. It, however, encounters difficulties when sub-categorizing 360-degree music videos because they can be either fictional or non-fictional. Therefore, Railton and Watson's (2011) classification of music videos was also adopted in this study. The two authors classified music videos into four categories: fictional music videos, music videos based on a true story, music performance videos, and documentary music videos. While the first two categories can be classified as fiction, the two latter can be classified as non-fiction.

b) Immersion: Immersion and interactivity were often considered together in prior VR video studies. Juul (2018) found that viewers could easily see themselves switching from traditional music videos to immersive music videos because the immersion and the interactivity made them feel "mesmerized" and amplified their emotions (p. 44). Pavlik (2021) examined 15 different immersive films from the *2020 Tribeca Film Festival Cinema360* using a VR headset. His analytical framework included six dimensions – interactivity, immersion, multisensory engagement, use of AI/data, use of first or third-person perspective, and use of natural user interface. As a result, the author found that high levels of interactivity and immersion made the

films highly enjoyable and “inspiring” (p. 9), while very little interactivity could make the film lackluster (p. 10). While media scholars tend to agree that immersion and interactivity are what draw viewers to immersive media (Gorisse, Christmann, Amato, & Richir, 2017; Juul, 2018; Sangra, 2019; Uskali et al., 2021; Goutier, de Haan, de Bruin, Lecheler, & Kruikemeier, 2021; Pavlik, 2021; Zhang, 2022), very few studies made an effort to study each concept separately. Our current study wants to fill that gap.

By definition, immersion refers to “a psychological state of an individual who feels absorbed into a VE [virtual environment] induced by specific stimuli” (Elmezeny et al., 2018). Hebbel-Seeger (2017) found that using a 360-degree camera during a sporting event can be a great supplement to a live broadcast. By placing a 360-degree camera in one of the front rows of the audience, viewers would be able to experience proximity to the game and to the audience around them, immersing them in the experience as if they were already there (p. 9). In brief, former research often addressed immersion as the feeling of being present in a virtual world (Sheikh, Brown, Evans, & Watson, 2016; Jones, 2017; Hebbel-Seeger, 2017; Elmezeny et al., 2018). Along with this definition, prior studies pointed out two dimensions of immersion in 360-degree VR videos: technical immersion and narrative immersion (Ryan, 2015; Sheikh et al., 2016; Elmezeny et al., 2018). Technical immersion refers to the visual and audio cues added to 360-degree videos to direct the viewer’s gaze and to follow the storyline. This adds to the viewer feeling fully present in the virtual world. Narrative immersion refers to story elements inside 360-degree videos that influence the viewer's experience. In these studies, the operational definition of technical immersion is clear. However, the definition of narrative immersion is not due to its incorporation of multi-dimensions. The only clear operational definition of narrative immersion is the spatial and temporal context, or information about location and time. Indeed,

based on Ryan (2015)'s explanation of the four dimensions of narrative immersion, we found the term narrative immersion very similar to storytelling richness, another key characteristic of VR videos to be addressed in this study.

In this section, it is more imperative to delve deeper into the multitude of effects of technical immersion than narrative immersion. On one hand, visual and audio cues can enhance the feeling of being present and guide viewers through the scenes. On the other hand, technical errors can also cause disruption to the viewer's feelings of immersion (Sheikh et al., 2016; Elmezeny et al., 2018). For the present study, we specifically looked for elements that enhance and elements that disrupt the feeling of being present.

i) Enhanced Technical Immersion: Among elements that help enhance technical immersion are visual cues, audio cues, and spatial audio (Sheikh et al., 2016; Argyriou, Economou, Bouki, & Doumanis, 2016; Elmezeny et al., 2018; Burt, 2019; Feng, Xie, & Lou, 2019; Argyriou, Economou, & Bouki, 2020; Bosman, Buruk, Jørgense, & Hamari, 2023). In several experimental studies, spatial audio has often been thought of as an emerging technology that directly influences users' visual attention and enhances their immersive experience by creating a realistic and interactive environment (Holm, Väänänen, & Battah, 2019; Hirway, Qiao, & Murray, 2022; Bosman et al., 2023). Therefore, in this study, while examining the effect of enhanced technical immersion as a whole, we also examined the effect of visual cues and spatial audio separately.

ii) Disruptive Technical Immersion: Technical elements that might hinder immersive effects and wipe out the illusion of being present in the virtual world include all kinds of real-world distractions and negative cues inside the immersive experience. These include stereoscopic misalignment, image distortions, occlusion conflicts, and even the inclusion of text

or human actors (Ijsselsteijn, de Ridder, Freeman, & Avons, 2000; Jones, 2017; Elmezeny et al., 2018). We also observed the visibility of tripods and tripod shadows, distracting background noise, drones and their shadows/reflections, and especially blurred areas caused by the wind in 360-degree videos. All these disruptive elements were considered in our content analysis of 360-degree videos.

c) Interactivity: By definition, interactivity refers to the active control that users are granted over media content (Sundar, Kang, & Oprean, 2017; Elmezeny et al., 2018; Pimentel, Kalyanaraman, Lee, & Halan, 2021; Zhang, 2022). Because interactivity is perceived as crucial in generating greater immersion, both factors have often been addressed together. In 360-degree VR video studies, the degree of user control has been perceived as being limited to just browsing, tilting, and rotating to look around (Jones, 2017) with "little interaction and technical inclusion possibilities" (de Bruin, de Haan, Kruikemeier, Lecheler, & Goutier, 2022). Very few studies attempted to classify the interactivity levels of 360-degree videos. Pimentel et al. (2021) proposed just two levels of interactivity: low and high. At the low level, there are regular, linear 360-degree videos. To be considered at a high level of interactivity, the videos must be dissected into "a series of vignettes from which the user can view in a non-linear fashion via user inputs" (Pimentel et al., 2021). In a more recent study, Zhang (2022) found that higher levels of interactivity made it easier and more concise for viewers to learn difficult concepts (p. 68). The author classified interactivity in immersive videos into three levels. At the low level, users can only choose where to look. At the medium level, users can interact with components to learn difficult scientific concepts. At a high level, users can choose to initiate many actions by using a VR controller. For example, these actions include inspecting cell information, playing audio instructions, and performing an action to get rid of bacteria. Although Zhang's (2022)

classification is more sophisticated, it is not practical and not completely suitable for studying the current YouTube VR platform. Therefore, we adopted and modified Pimentel et al.'s (2021) approach. At the low level, users can just look around the spherical video, pause, and resume viewing. At the high level, users can select a chapter of the video to view in a non-linear fashion, click on the card links at the end of the video to view more, choose to view different layers of images/videos inside the primary 360-degree video, or do all of the above. This approach did not account for the non-physical interaction users may perform in response to audio and visual cues in a 360-degree video.

d) Storytelling richness: Prior studies in the field of journalism and digital advertising suggested that incorporating a clear narrative in 360-degree VR videos can help stimulate and sustain emotion, create fluid storylines, and lead to more engaging experiences for viewers (Dolan & Perets, 2016; Jones, 2017; Feng et al., 2019). In an educational setting, using narratives, or stories, in producing VR experiences also makes it easier for learners to discern information and knowledge (Zhang, 2022; Petersen, Petkakis, & Makransky, 2022). Although there is an overall agreement on the positive impacts of storytelling in VR video production, the level or richness of storytelling is what was not easy to categorize. Previously, scholars measured the degree of narrative structure in VR video ads based on a 7-point scale for five storytelling-related questions (Feng et al., 2019). Or, they classified storytelling richness in VR experiences based on elements of storytelling. These elements include plot, character, dialogue, conflict, resolution, and point of view (Zhang, 2022). After computing the average scores for five items on the narrative structure degree survey, Feng et al. (2019) classified all video ads into low, moderate, and high levels. Separately, Zhang (2022) classified storytelling richness into three levels: minimal, basic, and advanced, each with detailed descriptions. At the minimal level, the

immersive experience has just one speaking character who delivers information through voiceover (VO) without any dialogue. At the basic level, the immersive experience has three speaking supporting characters who communicate with each other while the main character (the viewer) stays silent. At the advanced level, the immersive experience has a protagonist (the viewer), an antagonist, and at least four supporting characters with rich dialogue. In the current study, we modified Zhang's (2022) classification to include four levels, with the first level having no speaking characters (None), the second level having one speaking character with VO (Minimal), the third level having a non-drama dialogue between at least two characters (Basic), and the fourth level having a full dialogue between the protagonist, an antagonist, and their supporting characters (Advanced).

B. Audience Engagement

The practices of viewing, liking, and commenting on YouTube are most often considered "active audience engagement" under the Uses and Gratifications (U&G) framework. This a broad framework that classifies various user motivations in participating/using a media platform (Khan, 2017) and gratifications sought (Cheng, Yuan, & Wen, 2022). Arguably, by looking at audience metrics, researchers can see through the unconscious motivations that drive the audience to engage with a specific piece of content (Khan, 2017; Xie et al., 2022). This narrow behavioral aspect of audience engagement is currently attracting more attention from both content creators and scholars who want to generate insights from audience metrics (Manovic, 2018; Ferrer-Conill & Tandoc, 2018; Zamith, Belair-Gagnon, & Lewis, 2019; Steensen et al., 2020). The focus on audience metrics has been so prominent that some scholars put it as one of the twenty propositions of a new theory that depicts four dimensions and five relations of

audience engagement (Steensen et al., 2020). From a technical-behavioral dimension, the human-to-content aspect of audience engagement was explained as how users interact with written, oral, and visual texts (Steensen et al., 2020, p. 1668). Although there were some studies about audience engagement on the YouTube platform (Haridakis & Hanson, 2009; Khan, 2017; Wang et al., 2018; Xie et al., 2022), audience engagement with immersive videos on YouTube, especially non-journalism genres, has not been focused on enough. This study fills the gap in media research by studying how the audience interacted with VR videos on YouTube and quietly expressed their motivations sought through the number of views, likes, and comments.

C. Research Questions and Hypotheses

While the technical behavioral audience engagement theoretical framework allows for inspecting how viewers were attracted to YouTube VR videos, prior qualitative/experimental studies on key characteristics of VR allow us to come up with the following research questions and hypotheses.

RQ1: How were genre, immersion, interactivity, and storytelling richness featured on the YouTube VR channel?

RQ2: How did viewers react to VR videos on the YouTube VR channel?

H1: There is a statistically significant difference in audience engagement with respect to genre.

H2a: There is a statistically significant difference in audience engagement with respect to the use of enhanced technical elements.

H2b: There is a statistically significant difference in audience engagement with respect to the use of visual cues.

H2c: There is a statistically significant difference in audience engagement with respect to the use of spatial audio.

H2d: There is a statistically significant difference in audience engagement with respect to the existence of disruptive technical elements.

H2e: There is a statistically significant difference in audience engagement with respect to narrative immersion.

H3: There is a statistically significant difference in audience engagement with respect to interactivity.

H4: There is a statistically significant difference in audience engagement with respect to storytelling richness.

III. Methods

As of January 23, 2023, the YouTube VR channel at https://www.youtube.com/vr had 137 playlists and 1,523 VR videos. We retrieved these videos via YouTube Data API. After removing private, duplicated, non-VR videos, a random selection of 300 videos out of a list of 1,323 videos, or about 23% of the population, were selected for a quantitative content analysis using a pair of Oculus Quest 2 VR headsets. The sampling and coding process was guided by Riffe, Lacy, Fico, and Watson (2019).

For audience metrics, we collected the number of views, likes, and comments for each video in the sample using the two R packages *tuber* and *purr*. After that, analyses of covariance (ANCOVA) were run to compare the means of each audience engagement type (likes, comments, shares) across key characteristics (genre, immersion, interactivity, and storytelling richness),

controlling for the number of days the videos are on YouTube. The number of days was calculated based on the publishing date and the date all videos were collected (January 23, 2023).

Two coders first coded 30 videos (10% of the sample), and the intercoder reliability was 0.87 for genre, 0.86 for enhanced technical immersion, 0.86 for visual cues, 0.8 for disruptive technical immersion, 0.93 for narrative immersion, 0.91 for interactivity, and 0.83 for storytelling richness using Krippendorff's Alpha. As only one coder can detect spatial audio, we didn't calculate the intercoder reliability for this variable.

IV. Results

A. RQ1: How were key characteristics of VR videos featured on the YouTube VR channel?

a) Genre: Table 1 shows that 76% of VR videos in our sample are non-fiction ($N = 228$). About half of these videos ($N = 120$) are documentaries, news, and features produced by news organizations, non-profit/non-government organizations, and individual producers. Sporting ($N = 35$) and music ($N = 24$) are also popular genres, followed by cultural tours ($N = 23$). Other genres include vlogs, behind the scenes, cooking shows, fashion shows, and promotional videos.

Among fictional videos, animation is the most popular genre ($N = 28$), followed by film ($N = 22$), video game footage ($N = 12$), and music ($N = 10$). In total, fictional videos contribute to 24% of videos in our sample.

Table 1			
VR Genres			
Genre		**N**	**% of Total**
Fiction	Film	22	7.33%
	Animation	28	9.33%
	Video Game Footage	12	4.00%
	Music	10	3.33%
Non-Fiction	Documentaries	76	25.33%
	News and Features	44	14.67%
	Cultural Tour	23	7.67%
	Music	24	8.00%
	V-log	10	3.33%
	Athletic/Sporting Activity	35	11.67%
	Others	16	5.33%
Total		300	100.00%

b) Immersion: Table 2a shows that two-thirds of the videos had visual and/or audio cues to enhance viewers' feelings of immersion ($N = 200$). These cues included text, icons, animations, photos, videos, 360-degree videos, data visualizations, 3D objects, and visual effects.

Table 2a		
Enhanced Technical Immersion		
	N	**% of Total**
No	100	33.33%
Yes	200	66.67%
Total	300	100.00%

Table 2b shows that more than half of the videos had visual cues to enhance viewers' feelings of immersion ($N = 159$).

Table 2b		
Technical Immersion – Visual Cues Only		
	N	**% of Total**
No	141	47.00%
Yes	159	53.00%
Total	300	100.00%

Table 2c shows that only 12.67% of the videos had spatial audio ($N = 38$).

Table 2c		
Spatial Audio		
	N	**% of Total**
No	262	87.33%
Yes	38	12.67%
Total	300	100.00%

Table 2d presents the fact that more than half of the videos had disruptive elements that weaken the feeling of immersion ($N = 156$). The most prevalent issues are visible stitching lines, camera tripods, camera shadows and reflections, and other video editing mistakes. Issues can also be caused by distorted characters or images due to close proximity to the camera and/or distracting background noise.

Table 2d		
Disruptive Technical Immersion		
	N	**% of Total**
No	144	48.00%
Yes	156	52.00%
Total	300	100.00%

Table 2e shows that two-thirds of the videos provided context and background information, allowing the viewers to feel more immersed ($N = 203$).

Table 2e		
Narrative Immersion		
	N	**% of Total**
No	97	32.33%
Yes	203	67.67%
Total	300	100.00%

c) Interactivity: Due to the limited features of YouTube, nearly 70% of videos only allowed viewers to choose where to look inside an experience ($N = 209$) (See Table 3). Just about 30% of the videos allowed viewers to either jump to a specific segment within an experience, view extra layers of information, click on a card link, or follow a call-to-action message. These call-to-action messages often include calls for subscribing, liking, or joining a cause.

Table 3		
Level of Interactivity		
Level	**N**	**% of Total**
Low	209	69.67%
High	91	30.33%
Total	300	100.00%

d) Storytelling Richness: It's worth noting that 21.33% of the videos did not have a story plot to begin with ($N = 64$) (See Table 4). They did not have any speaking characters, nor did they have non-speaking characters. These videos more often just had background music. Table 3 shows that 46% of the videos had just one speaking character ($N = 138$) or were classified at a minimal level of storytelling richness. One-fifth of the videos had a basic level of storytelling richness ($N = 64$), which is classified as having at least two speaking characters with dialogue. Just over 11% of the videos had a protagonist, an antagonist, and supporting characters with dialogue ($N = 34$), fitting in the advanced level of storytelling richness.

Table 4		
Level of Storytelling Richness		
Level	**N**	**% of Total**
None	64	21.33%
Minimal	138	46.00%
Basic	64	21.33%
Advanced	34	11.33%
Total	300	100.00%

e) VR 180 Format: An unexpected finding is that about 7% of the videos were produced in the VR 180-degree format ($N = 21$) (See Table 5). This format is an improved version of traditional 2D videos, but they may make viewers lose the feeling of immersion.

Table 5		
VR Formats		
	N	**% of Total**
VR 360	279	93.00%
VR 180	21	7.00%
Total	300	100.00%

B. RQ2: How do viewers react to VR videos on the YouTube VR channel?

a) Audience Engagement: Regarding audience engagement, Table 6 shows that on average, each video has 1.7 million views, 11,336 likes, and 310 comments. Standard deviations are huge across all types of engagement, $SD = 1{,}711{,}972.93$, $SD = 11{,}336.77$, and $SD = 310.3$, respectively, indicating the skewness of our sample data. On average, each video lasts about 288.11 seconds, or 4.8 minutes, and has been on YouTube for about 1,299.01 days, or 3.6 years. Standard deviations for these two variables are fairly normal, $SD = 288.11$ and $SD = 824.51$, respectively.

Table 6					
Descriptive Statistics of Audience Engagement, Video Length, and Days on YouTube (N = 300)					
	Views	**Likes**	**Comments**	**Video Length (in seconds)**	**Days on YouTube**
Mean	1,711,972.93	11,336.77	310.34	288.11	1,299.01
SD	10,208,576.41	28,639.92	909.07	236.43	824.51
Min	99.00	19.80	0.00	6.00	10.00
Max	159,066,911.00	669,071.00	8,972.00	1,802.08	2,863.00
Median	128,899.00	186.00	35.00	227.50	1,441.00

b) Analysis of Covariance: The ANCOVA tests were used to detect the impacts of independent variables taking into consideration a covariate. In this case, we used days on YouTube as a covariate and conducted an ANCOVA to test each of our eight hypotheses.

H1: There is a statistically significant difference in audience engagement with respect to genre. While controlling for the number of days on YouTube, the genre has a statistically significant impact on views, likes, and comments, $F(1, 297) = 4.608$, $p < .05$, $F(1, 297) = 8.54$, $p < .01$, and $F(1, 297) = 12.1$, $p < .001$, respectively. On average, fictional videos had more audience engagement than non-fictional videos.

H2a: There is a statistically significant difference in audience engagement with respect to the use of enhanced technical elements. Enhanced technical elements did not appear to have a statistically significant impact on audience engagement.

H2b: There is a statistically significant difference in audience engagement with respect to the use of visual cues. Visual cues did not appear to have a statistically significant impact on audience engagement.

H2c: There is a statistically significant difference in audience engagement with respect to the use of spatial audio. Spatial audio did not appear to have a statistically significant impact on audience engagement.

H2d: There is a statistically significant difference in audience engagement with respect to the existence of disruptive technical elements. While controlling for the number of days on YouTube, disruptive components have a statistically significant impact on views and likes, $F(1, 297) = 4.093$, $p < .05$ and $F(1, 297) = 5.602$, $p < .05$, respectively. Videos that do not have disruptive components tend to have more views and likes.

H2e: There is a statistically significant difference in audience engagement with respect to narrative immersion. While controlling for the number of days on YouTube, narrative immersion has a statistically significant impact on comments, $F(1, 297) = 6.249$, $p < .05$. However, in contrast to our H2e hypothesis, videos that have narrative immersion tended to have fewer comments.

H3: There is a statistically significant difference in audience engagement with respect to interactivity. We did not detect a significant difference in audience between interactive levels.

H4: There is a statistically significant difference in audience engagement with respect to storytelling richness. While controlling for the number of days on YouTube, storytelling richness has a statistically significant impact on likes and comments, $F(1, 295) = 2.841$, $p < .05$ and $F(1, 295) = 4.566$, $p < .01$, respectively.

It's worth noting that except for the number of views affected by technical disruptions, Levene's tests indicated unequal variances across other independent variables. As a result, just a portion of the H2d hypothesis was supported, specifically that videos tend to have more views if they do not have technical disruptions. When performing the Kruskal-Wallis tests on storytelling

richness, it was found that audience engagement (views, likes, comments) tends to vary depending on the level of richness. Post-hoc pair comparisons (Dunn's tests) indicated the difference between the advanced and minimal levels. As a result, a portion of our H4 hypothesis was supported, specifically that videos tend to have more views, likes, and comments if they have a richer story plot.

V. Discussion

Despite non-fiction videos making up almost 80% of our sample, fictional videos were the most popular category, consistent with Khan's (2017) highlighted entertainment motive for YouTube users. While non-fiction live, news and features, and documentaries are the most common types, fiction is a very popular genre that can reach a broad audience (Uskali et al., 2021). Fiction films and animation tend to have more characters and more immersive elements than other genres. This is due to the fact that the viewer often assumes the role of a character in the story. This gives the viewer the opportunity to interact with the video and with other characters, which increases engagement and immersion. Many fiction animation videos are also based on widely popular video games such as "Minecraft" or "Among Us." This connection to a well-known property makes it more likely that these videos will be found and engaged with.

Interactivity had no impact on audience engagement. 70% of the videos in the sample only allowed viewers to choose where to look. Just 30% of the videos provided some type of interactive component for the viewer within the video. While our hypothesis H3 was not supported, it is unsurprising considering the limitations of the YouTube platform. In previous studies that focused specifically on interactivity, platforms other than YouTube were used to host and view videos. This allowed the researchers to include more components that YouTube doesn't

allow, such as objects that viewers can physically interact with using a controller, or tasks that viewers must complete before the video can continue (Zhang, 2022). YouTube doesn't have these types of mechanics. The most interactivity it could offer would be cutting videos into segments and including clickable links/cards at the end. However, this form of interaction isn't significant or immersive enough to capture the viewer's attention. Viewers would have to be able to actually interact with components within the video and make changes to the story or environment to have any sort of significant impact, which is currently impossible on YouTube. However, while it was not observed in this study, there are popular 2D interactive video series on YouTube that could provide an example for more interactive VR videos in the future. These 2D series are "choose your own adventure" style videos that require viewers to make choices by clicking cards at the end of the video. This would then redirect them to other videos that influence the outcome of the story based on their choices. The most well-known series is a YouTube Original collaboration with Mark Fischbach titled, "In Space with Markiplier." However, it is unknown if these types of interactive series have been made for 360-degree videos, or if they would function as intended in any VR headset.

Storytelling richness had a statistically significant impact on likes and comments. Despite having the least amount of videos, the advanced category for storytelling richness had the highest average number of views, likes, and comments, which is consistent with earlier research. VR videos that have multiple speaking characters, including a protagonist and an antagonist, were found to be more positively received and fostered more engagement than those with fewer characters (Pavlik, 2021; Zhang, 2022).

While disruptive technical elements had a statistically significant impact on views and likes, beneficial technical elements were not statistically significant. Videos without disruptive

technical elements garnered many more views, likes, and comments than videos with disruptive elements. However, while videos with beneficial technical elements had more views, likes, and comments than videos without any beneficial cues, it wasn't statistically significant. This is fairly consistent with previous research. VR videos that have strong beneficial technical cues and no disruptive technical cues tend to be more immersive and foster more engagement (Elmezeny et al., 2018). This is especially true for videos that lack disruptive cues, as those videos garnered significantly more engagement than videos with disruptive cues. Another interesting note is the inclusion of spatial audio in some of the videos. Only 38 of the videos included spatial audio, and though it didn't have a statistically significant impact on audience engagement. However, the use of spatial audio is a relatively new and exciting feature. If it were implemented in more videos, it's possible that it could have a larger effect on audience engagement and interactivity, as voices in a 3D space could give explicit directions for a viewer to turn their head and view certain scenes. However, there are not enough videos with spatial audio in our study to fully grasp the effects it may have on audience engagement. Holm, Väänänen, and Battah (2019) also found that users were less interested in the active use of the spatial audio, although they gave the highest scores for listening to spatial audio in a 360-degree music setting.

Narrative immersion, despite having an effect, did not have a statistically significant impact on audience engagement. However, it did have the opposite effect on engagement than we originally hypothesized. Videos without a speaking character that provided spatial and temporal context tended to foster more audience engagement than videos that included this context. This contradicts previous research. Previous studies have found that narrative immersion can lead to a better viewing experience and foster more audience engagement (Elmezeny et al., 2018; Zhang, 2022). Despite this contradiction, it is not entirely unsurprising that these results

were different from what we hypothesized. One of the most common types of VR videos that lacked narrative immersion in our study were videos that fell under the category of "music and video to relax to." These are videos that only include calming background music and simple footage or still images. This is a fairly common type of 2D video on YouTube. They are often described as videos that help the viewer relax or study, and it is one of the more popular genres on YouTube.

VI. Conclusion

Our study contributes to media discourse on audience engagement with immersive media content by presenting more information about the behavioral aspect of viewer-to-content interactions (Khan, 2017; Wang et al., 2018; Steensen et al., 2020). Our results reveal that viewers are more likely to engage if the 360-degree videos have a more sophisticated level of technical immersion, a higher level of richness in storytelling, and to some extent, a fictional component. The findings have significant implications for VR video producers. A few production tips are suggested to help enhance the quality of immersive videos such as minimizing disruptive elements (distorted characters/objects, visible camera tripods, reflections of camera equipment, background noise, etc.) and increasing the level of richness in storytelling (having more speaking characters with significant roles in the story).

This research has limitations that need to be addressed in future studies. First, in this present study, we have not come up with multilevel scales for measuring immersion, interactivity, and storytelling richness. Second, comments from the viewers have not been analyzed to reveal viewer-to-viewer interactions as suggested in Steensen et al.'s (2020) audience engagement model. Third, the impacts of spatial audio should be examined on a larger

scale. Finally, our results can only be generalized for videos on the YouTube VR channel, not for videos on other YouTube accounts.

References

Argyriou, L., Economou, D., Bouki, V., & Doumanis, I. (2016). Engaging Immersive Video Consumers: Challenges Regarding 360-Degree Gamified Video Applications. 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS), 145–152. https://doi.org/10.1109/IUCC-CSS.2016.028

Argyriou, L., Economou, D., & Bouki, V. (2020). Design methodology for 360° immersive video applications: the case study of a cultural heritage virtual tour. Personal and Ubiquitous Computing, 24(6), 843–859. https://doi.org/10.1007/s00779-020-01373-8

Bleumers, L., Van den Broeck, W., Lievens, B., & Pierson, J. (2012). Seeing the bigger picture: a user perspective on 360 TV. In: Proceedings of the 10th European Conference on Interactive TV and Video, pp. 115–124. ACM

Bosman, I., Buruk, O., Jørgensen, K., & Hamari, J. (2023). The effect of audio on the experience in virtual reality: a scoping review. *Behaviour & Information Technology*. 10.1080/0144929X.2022.2158371. (1-35).

Burt, M. (2019). *What do audiences want from a virtual reality experience?* (thesis). *EPrints*. https://eprints.qut.edu.au/135430/1/Malcolm_Burt_Thesis.pdf

Cheng, Y., Yuan, W., & Wen, Z., (2022). "Shared Virtual Reality Experiences during the COVID-19 Pandemic: Exploring the Gratifications and Effects of Engagement with

Immersive *Videos" International Journal of Environmental Research and Public Health* 19, no. 9: 5056. https://doi.org/10.3390/ijerph19095056

Cummings, J. J., Tsay-Vogel, M., Cahill, T. J., & Zhang, L. (2022). Effects of immersive storytelling on affective, cognitive, and associative empathy: The mediating role of presence. *New Media & Society*, 24(9), 2003–2026. https://doi.org/10.1177/1461444820986816

de Bruin, K., de Haan, Y., Kruikemeier, S., Lecheler, S., & Goutier, N. (2022). A first-person promise? A content-analysis of immersive journalistic productions. *Journalism* (London, England), *23*(2), 479–498. https://doi.org/10.1177/1464884920922006

Elmezeny, A., Edenhofer, N., & Wimmer, J. (2018). Immersive storytelling in 360-degree videos: An analysis of interplay between narrative and technical immersion. *Journal For Virtual Worlds Research*, *11*(1). https://doi.org/10.4101/jvwr.v11i1.7298

Ekström, M. (2002). Epistemologies of TV journalism: A theoretical framework. *Journalism* (London, England), 3(3), 259–282. https://doi.org/10.1177/146488490200300301

Feng, Y., Xie, Q., & Lou, C. (2019). The Key to 360-Degree Video Advertising: An Examination of the Degree of Narrative Structure. *Journal of Advertising*, *48*(2), 137–152. https://doi.org/10.1080/00913367.2019.1585305

Ferrer-Conill, R., & Tandoc Jr., E. C. (2018). The Audience-Oriented Editor. *Digital Journalism* 6: 436–453. doi: 10.1080/21670811.2018.1440972

Feurstein, M. S. (2018). Towards an Integration of 360-Degree Video in Higher Education. In Daniel Schiffner (Ed.), *Proceedings of DeLFI Workshops 2018 co-located with 16th e-Learning Conference of the German Computer Society (DeLFI 2018)* (pp. 1 - 12). CEUR WS. http://ceur-ws.org/Vol-2250/WS_VRAR_paper3.pdf

Foxman, M., Beyea, D., Leith, A. P., Ratan, R. A., Chen, V. H. H., & Klebig, B. (2022). Beyond Genre: Classifying Virtual Reality Experiences. *IEEE Transactions on Games*, *14*(3), 1–1. https://doi.org/10.1109/TG.2021.3119521

Godulla, A., Planer, R., Wolf, C., Lück, A., & Vaaßen, F. (2021). An Immersive Journey through Flawed Technology: Users' Perceptions of VR in Journalism. *Journalism and Media*, *2*(3), 454–468. https://doi.org/10.3390/journalmedia2030027

Gorisse, G., Christmann, O., Amato, E. A., & Richir, S. (2017). First- and Third-Person Perspectives in Immersive Virtual Environments: Presence and Performance Analysis of Embodied Users. *Frontiers in Robotics and AI*, *4*(Article 33), 12–. https://doi.org/10.3389/frobt.2017.00033

Goutier, N., de Haan, Y., de Bruin, K., Lecheler, S., & Kruikemeier, S. (2021). From "Cool observer" to "emotional participant": The practice of immersive journalism. *Journalism Studies*, *22*(12), 1648–1664. https://doi.org/10.1080/1461670x.2021.1956364

Haridakis, P., & Hanson, G. (2009). Social Interaction and Co-Viewing With YouTube: Blending Mass Communication Reception and Social Connection. *Journal of Broadcasting & Electronic Media*, *53*(2), 317–335. https://doi.org/10.1080/08838150902908270

Hebbel-Seeger, A. (2017). 360 degrees video and VR for training and marketing within sports. *Athens Journal of Sports*, *4*(4), 243–262. https://doi.org/10.30958/ajspo.4.4.1

Hirway, A., Qiao, Y., & Murray, N. (2022). Spatial audio in 360° videos: does it influence visual attention? In proceedings of the 13th ACM Multimedia Systems Conference (MMSys '22). *Association for Computing Machinery*, New York, NY, USA, 39–51. https://doi.org/10.1145/3524273.3528179

Holm, J., Väänänen, K., & Battah, A. (2020). User experience of stereo and spatial audio in 360° live music videos. *Proceedings of the 23rd International Conference on Academic Mindtrek*, 134–141. https://doi.org/10.1145/3377290.3377291

Ijsselsteijn, W. A., de Ridder, H., Freeman, J., & Avons, S. E. (2000). Presence: Concept, determinants, and measurement. In Proceedings, Human Vision and Electronic Imaging V, 3959, 520-529. doi: 10.1117/12.387188

Jones, S. (2017). Disrupting the narrative: immersive journalism in virtual reality. Journal of Media Practice, 18(2-3), 171–185. https://doi.org/10.1080/14682753.2017.1374677

Johnson, J. (2019). Jumping into the World of Virtual and Augmented Reality. *Knowledge Quest*, *47*(4), 22–27. https://eric.ed.gov/?id=EJ1207783

Juul, L. (2018). Examining 360° storytelling in immersive music videos (Dissertation, Malmö universitet/Teknik och samhälle). http://urn.kb.se/resolve?urn=urn:nbn:se:mau:diva-20828

Kim, Y. & Lee, H. (2022). User acceptance of 360-degree video news: an integrated model of extended TAM and U&G perspectives. *Communication Research and Practice*, ahead-of-print(ahead-of-print), 1–20. https://doi.org/10.1080/22041451.2022.2141861

Maas, M. J., & Hughes, J. M. (2020). Virtual, augmented and mixed reality in K–12 education: A review of the literature. *Technology, Pedagogy and Education*, *29*(2), 231–249. https://doi.org/10.1080/1475939x.2020.1737210

Manovich, L. (2018). Digital Traces in Context| 100 Billion Data Rows per Second: Media Analytics in the Early 21st Century. *International Journal of Communication, 12*, 473–488. https://ijoc.org/index.php/ijoc/article/view/6160/2249

Pavlik, J. V. (2021). Experiencing Cinematic VR: Where Theory and Practice Converge in the Tribeca Film Festival Cinema360. *Proceedings of the New York State Communication Association*, *2020*(9). https://docs.rwu.edu/nyscaproceedings/vol2020/iss1/9/

Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers and Education*, 179, 104429–. https://doi.org/10.1016/j.compedu.2021.104429

Pimentel, D., Kalyanaraman, S., Lee, Y. H., & Halan, S. (2021). Voices of the unsung: The role of social presence and interactivity in building empathy in 360 video. *New Media & Society*, 23(8), 2230–2254. https://doi.org/10.1177/1461444821993124

Prothero, J. D., & Hoffman, H. D. (1995). Widening the field-of-view increases the sense of presence within immersive virtual environments. *Human Interface Technology Laboratory Technical Report R-95-4*. Seattle: University of Washington.

Railton, D. & Watson, P. (2011). *Genre and Music Video: Configurations and Functions*. 10.3366/edinburgh/9780748633227.003.0003.

Riffe, D., Lacy, S., & Fico, F. (2019). *Analyzing media messages : using quantitative content analysis in research* (4th edition.). Lawrence Erlbaum. https://doi.org/10.4324/9780429464287

Ryan, M. L. (2015). *Narrative as virtual reality 2: Revisiting immersion and interactivity in literature and electronic media*. Baltimore: John Hopkins University Press. https://doi.org/10.1353/book.72246

Sangra, S. (2019). Scope and Ethical Implications of Immersive Storytelling and Journalism. *Global Media Journal*. https://amityuniversity.ae/gmj-ae/journals/Seema-Sangra.pdf

Sheikh, A., Brown, A., Evans, M., & Watson, Z. (2016). Directing attention in 360-degree video. Paper presented at *IBC 2016 Conference,* Amsterdam. doi: 10.1049/ibc.2016.0029

Steensen, S., Ferrer-Conill, R., & Peters, C. (2020). (Against a) Theory of Audience Engagement with News. *Journalism Studies* (London, England), 21(12), 1662–1680. https://doi.org/10.1080/1461670X.2020.1788414

Sundar, S., Kang, J., & Oprean, D. (2017). Being there in the midst of the story: how immersive journalism affects our perceptions and cognitions. *Cyberpsychology, Behavior, and Social Networking*, *20*(11): 672–682. https://doi.org/10.1089/cyber.2017.0271

Tricart, C. (2018). *Virtual reality filmmaking : techniques and best practices for VR filmmakers* (1st edition). Taylor and Francis. https://doi.org/10.4324/9781315280417

Uskali, T., Gynnild, A., Jones, S., & Sirkkunen, E. (2021). *Immersive Journalism as Storytelling: Ethics, Production and Design*. *Oapen*. Routledge. https://library.oapen.org/handle/20.500.12657/42893

Wang, G., Gu, W., & Suh, A. (2018). The Effects of 360-Degree VR Videos on Audience Engagement: Evidence from the New York Times. *HCI in Business, Government, and Organizations*, 217–235. https://doi.org/10.1007/978-3-319-91716-0_17

Xie, W., Damiano, A., & Jong, C. H. (2022). Emotional Appeals and Social Support in Organizational YouTube Videos During COVID-19. *Telematics and Informatics Reports*, 100028. https://doi.org/10.1016%2Fj.teler.2022.100028

Young, M. L., Hermida, A., & Fulda, J. (2018). What Makes for Great Data Journalism? *Journalism Practice, 12*(1), 115-135. https://doi.org/10.1080/17512786.2016.1270171

Zamith, R., Belair-Gagnon, V., & Lewis, S. C. (2019). Constructing Audience Quantification: Social Influences and the Development of Norms About Audience Analytics and Metrics. *New Media & Society*. doi: 10.1177/1461444819881735

Zhang, L. (2022). *Investigating interactivity and storytelling in immersive virtual reality for science education* (dissertation). Virginia Tech, Blacksburg, Virginia. https://vtechworks.lib.vt.edu/handle/10919/107808