

Behind the Curtain of the “Ultimate Empathy Machine”: On the Composition of Virtual Reality Nonfiction Experiences

Chris Bevan
University of Bristol
Bristol, UK
chris.bevan@bristol.ac.uk

David Phillip Green
University of the West of England
Bristol, UK

Harry Farmer
University of Bath
Bath, UK

Mandy Rose
University of the West of England
Bristol, UK

Kirsten Cater
University of Bristol
Bristol, UK

Danaë Stanton Fraser
University of Bath
Bath, UK

Helen Brown
University of Bath
Bath, UK

ABSTRACT

Virtual Reality nonfiction (VRNF) is an emerging form of immersive media experience created for consumption using panoramic “Virtual Reality” headsets. VRNF promises non-fiction content producers the potential to create new ways for audiences to experience “the real”; allowing viewers to transition from passive spectators to active participants. Our current project is exploring VRNF through a series of ethnographic and experimental studies. In order to document the content available, we embarked on an analysis of VR documentaries produced to date. In this paper, we present an analysis of a representative sample of 150 VRNF titles released between 2012-2018. We identify and quantify 64 characteristics of the medium over this period, discuss how producers are exploiting the affordances of VR, and shed light on new audience roles. Our findings provide insight into the current state of the art in VRNF and provide a digital resource for other researchers in this area.

CCS CONCEPTS

• **Human-centered computing** → **Virtual reality**; • **Applied computing** → **Media arts**; • **Software and its engineering** → *Virtual worlds software*.

KEYWORDS

Virtual Reality, nonfiction, immersive media, interaction.

ACM Reference Format:

Chris Bevan, David Phillip Green, Harry Farmer, Mandy Rose, Kirsten Cater, Danaë Stanton Fraser, and Helen Brown. 2019. Behind the Curtain of the “Ultimate Empathy Machine”: On the Composition of Virtual Reality Nonfiction Experiences. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019)*, May 4–9, 2019, Glasgow, Scotland UK. ACM, New York, NY, USA, 12 pages. <https://doi.org/10.1145/3290605.3300736>

1 INTRODUCTION

Virtual Reality nonfiction (VRNF) is a relatively recent form of immersive interactive media experience that is created for consumption using VR head mounted displays (VRHMD). VRNF content, often categorised in app stores using terms such as ‘thought provoking experiences’ has a broad thematic range, including documentary, journalism, history and exploration.

The promise of VRNF is to push beyond film and invite viewers to immerse themselves in different lives and worlds; to transition from being passive spectators to becoming active participants. This potential has attracted impressive claims, most notably that nonfiction content presented in VR might be the “ultimate machine for empathy” [30].

Of particular relevance to the HCI community, the production of VRNF is a multidisciplinary practice with - as yet - no



Figure 1: *Hunger In Los Angeles* (2012). Viewer on right is using prototypical VR hardware that would later become the Oculus Rift.

established independent design grammar [6]. VRNF experiences are drawing inspiration (and blending characteristics) from gaming and immersive theatre, alongside traditional filmmaking practices [22]. However, despite growing content and exposure within the VR and filmmaking communities, VRNF is relatively understudied by the HCI community.

VRNF is an emergent medium and its distribution is fragmented. As such it is currently very difficult to gain a birds-eye view of the range of content available and how it is evolving. For example, some pieces are self-funded by artists and creative communities, distributed only through film festivals and exhibition, while others are produced by specialist VRNF studios (e.g. *Felix & Paul Studios*, *Emblematic*, *Here Be Dragons*) and distributed through a range of online content portals. Elsewhere, international news organisations including *Guardian*, *New York Times*, *BBC* and *Al Jazeera* have experimented with VRNF, distributing content through their own online portals.

Our project is bringing together a multidisciplinary team of researchers from HCI, psychology and documentary studies to explore VRNF through a series of ethnographic and experimental studies. To support this investigation, we have embarked on the creation of a unique catalogue of VRNF media content, a *Mediography of Virtual Reality Nonfiction*, which will be released as a publicly available online research tool. Thus far, we have identified and indexed 350 titles, with associated title metadata including director, year of release and duration.

In this paper, we respond to the need for a detailed examination of this new form of interactive media with the in-depth analysis of a large representative sample of VRNF taken from our repository, covering much of the first six years of its history. We consider our contributions to the research community to be fourfold:

- (1) **Insights on the internal composition of VRNF** as new form of interactive media experience.
- (2) **The identification and quantification of key characteristics** that describe the composition of VR non-fiction content.
- (3) **Insights from an examination of the role of the viewer in VRNF**, and how producers of VRNF are creating content to support new types of viewer role.
- (4) **The production of a unique open-access VRNF research tool** with search and filtering capabilities enhanced by the additional metadata generated from point 2.

A Brief History of VRNF

VRNF reflects two histories - VR (virtual reality) + NF (non-fiction). The recent intersection of these histories forms the focus of our research. It has come at an intriguing historical moment, as society grapples with issues such as veracity, authenticity, trust, and the role of media technologies in the 'post truth' era. We contend that VRNF has great potential to provide authentic, complex, legible and enjoyable experiences, grounded in truth, that present new ways of engaging with and making sense of the world. Various interaction design challenges are suggested by this intersection. To articulate (and ultimately address) these challenges, it is important to begin by acknowledging the individual histories involved.

A Brief History of VR. Virtual Reality has a rich history that draws on numerous technological innovations, including Wheatstone's *Stereoscope* (1838), and Ader's *Theatrophone* (1881). Sutherland's *Sword of Damocles* (1968), widely considered to be the first VRHMD, bears a striking resemblance to more recent HMDs such as *Oculus Go* (2018). Yet their visual similarities belie innovations in specification and UX, as well as nascent cultures of content production and distribution, that clearly distinguish them as the products of different era.

A recent perspective on the modern history of Virtual Reality [25] includes 52 definitions of VR, which explodes its concept beyond any specific apparatus. One key distinction is that VR tends to refer to fully-immersive experiences [29]. This is in contrast with mixed / augmented reality experiences that include elements of the viewer's physical environment. Yet even the most advanced VR technology does not filter 'real environments' entirely, so we must adopt such distinctions carefully. For the time being, we simply need to be specific about what we mean by "VR", and at the same time acknowledge its potential polysemy.

The affordances of VRHMDs has resulted in a popular lexicon of associated terms; "embodiment" [8]; "immersion" [31]; "presence" (conceptualised as a combination of 'place' and 'plausibility' illusions [39]); "empathy" (a term that inspires

fierce debate [5]). The allure of VR is that it supposedly lends itself naturally to these qualities, but we must be careful; many of these terms are themselves loosely-defined [10]. Some also question the hype surrounding VR [43] [36] and undoubtedly the technology has limitations. Yet producers invariably work creatively within - and often inspired by - limitations [23]. We are interested in what we can learn from the outputs of creative professionals producing content within the limitations of existing VR technology infrastructure; what can that tell us about VR? What can that tell us about nonfiction?

A Brief History of NF. Nonfiction media plays an important cultural and social role in most human societies [17]. It helps us make sense of the world by focusing our attention on a particular topic, telling stories using narrative language specific to the medium and techniques such as genre conventions. The role of technology in this process pre-dates the digital era [45] and spans print, radio, television and online/social media; VR is the latest addition. In the era of 'post truth', and as media technologies become increasingly powerful, it is more important than ever that we "look behind the curtain" of the technologies involved in nonfiction storytelling. We need to understand the technical and creative constructs upon which these stories are produced. Unfortunately, gaining a deep understanding of a particular medium is increasingly challenging. The well-documented, historical practices of nonfiction media-making (e.g. journalism and documentary) are constantly evolving and converging with a wide variety of other media-making practices [22] - social media, interactive formats (e.g. i-doc [2]) and elements such as ubiquitous computing [16]. Many of these newer practices are only poorly (or partially) understood [21]. The media grows ever more complex. As VRNF emerges into this complex, now is an opportune moment to adopt a systematic approach and start at the beginning.

Our starting point is 2012, when the American journalist and film-maker Nonny De La Peña began experimenting with a prototype VRHMD. De la Peña's work - later described as "immersive journalism" [11] - took audio recordings of real world events and added computer-generated imagery to provide immersive visual context. This process led to the development of the Oculus *DK1* and a resulting work - *Hunger in LA* (fig. 1) - became the first piece of VRNF to be exhibited at an international film festival.

Since then, VRNF has been steadily increasing in visibility at international festivals (e.g. Tribeca, Sundance and Venice) and there has been a surge of interest and investment. In January 2018, the VRNF title *Spheres* achieved worldwide press attention when it was acquired for a seven figure sum [44]. In parallel, consumer VRHMDs (e.g. Oculus *Rift* and HTC *Vive*), distribution platforms (e.g. *Steam*, *Oculus Store*),

and apps (e.g. *Jaunt*), have brought VRNF experiences to home consumers, and production tools such as 360 video cameras and games engines (e.g. *Unity*) have contributed to an ever-growing toolset for VRNF producers. But how are these tools being put to use? What does VRNF look like?

An oft-cited example of VRNF is *Clouds over Sidra* (2015). It tells the story of 13 year old Sidra, a resident of the Za'atari refugee camp in Jordan. In the piece, we experience Sidra's environment via a series of static 360 degree panoramic videos; Sidra's voiceover describes day to day life on the refugee camp. It was this work that inspired its producer Chris Milk to articulate the infamous conceptualisation of VR as the "ultimate empathy machine", and charity organisations were quick to explore the potential of VRNF as a driver for prosocial behaviour [3] (while others question the ethical basis for such assumptions [5]). A critically-acclaimed example of VRNF is *Notes on Blindness: Into Darkness* (2016) which (like *Hunger in LA*) illustrates archive audio recordings (this time taken from the audio diary John Hull, who lost his sight in his 40s) with evocative visuals that simulate Hull's descriptions of visual impairment. *Carne y Arena* (2017) is an ambitious, room-scale installation, which actually includes several 'rooms', requiring the viewer to navigate through real space and interact with actors. Typically, VRNF uses a combination of interactive elements and traditional visual storytelling techniques (video, animation, audio, editing, voiceover, etc) presented as a 360 degree panorama via a VRHMD.

More than the sum of its parts? As these examples demonstrate, some elements of VRNF (e.g. 360 video) clearly take their cue from earlier nonfiction genres such as documentary and journalism. Other elements (e.g. CGI virtual environments) draw upon other media such as video games, while others take inspiration from immersive theatre. This mashup of multiple influences calls for a ground-up approach, rather than one that derives from existing frameworks of analysing media.

2 RELATED WORK

To our knowledge, ours is the first attempt to identify, collate and analyse the characteristic features of VRNF media en masse. However, there are a number of related works that have motivated our work and inspired our approach. Gifreu-Castells [18] for example has examined key works and characteristics within the field of interactive documentary, and recent works by Rose [35] [36] and Nash [32] have contextualised VRNF within the broader canon of interactive documentary (and VR itself). Other researchers including Tran [42] are examining factors relating to the quality of experience of immersive VR video content, if not necessarily the composition of the content itself.

In terms of how VRNF is constructed, a multi-disciplinary group of researchers are working to identify and develop best practices across a range of HCI relevant topics that complement our work here. Such topics include subtitle placement and management [7], how viewers transition to and from VR [24], the role of sound [34], the legibility of text in VR environments [12] and interpersonal communication within embodied VR experiences [40].

More broadly, although this latest wave of VR hardware is very new, research into the qualities of immersive 3D virtual environments for digital storytelling and its potential benefits for areas such as education and social justice are extensive and likely instructive. The works of Schuemie [37], Mikropoulos [28], McQuiggan [27] and Robin [33] provide an excellent overview of the state of the art prior to this latest iteration of Virtual Reality hardware. Similarly, interactive documentary itself has a long history, with extensive works from institutions such as the Digital Storytelling Lab at Columbia, the Open Documentary Lab at MIT and the UK i-Docs project among others.

Despite what we have learned from desktop 3D immersive environments however, the degree to which this new wave of Virtual Reality effects is capable of evoking new levels of emotional engagement and response to nonfiction content - to effect its promise as “the ultimate machine for empathy” - is far from clear. What we can say however is that this latest iteration of VR hardware is the first commercially available form of VR that is able to achieve the “representational fidelity” (e.g. smooth rendering of real time photorealistic graphics in an HMD) that Dalgarno and Lee [9] argued as being crucial for facilitating a sense of presence in a virtual world. The challenge of VRNF nonfiction content producers therefore (and a major motivation for this work) is recognising that the ability to provide an immersive 360 degree environment may not itself be sufficient to make a viewer feel present in a story [38]. Facilitating a sense of ‘being there’ in VRNF calls for new interaction design challenges that we hope this work will enable future work to overcome.

3 THE CHARACTERISTICS OF VRNF

In this section, we describe the selection and analysis of a large representative sample of English language VRNF produced between 2012 and the first quarter of 2018. The objectives of our analysis were twofold: 1) to identify and quantify a set of generalisable features of VRNF media, and 2) to examine the role of the viewer within VRNF experiences, and how content creators are utilising VR to explore new viewer roles.

We recognise that what constitutes true “VR” and what is considered “nonfiction” remains a source of contention across disciplines (e.g. [14], [18]). We are aware for example

that there is a strong line of reasoning within computer science and HCI that rejects 360 degree video as a valid example of VR, due to it lacking any meaningful form of simulation [26]. Intuitively, we can also argue that the affordance of CGI in VR to reconstruct and simulate real-world people, places and events inevitably blurs the relationship between what is acceptable as fact or fiction. Resolving these arguments is not within the scope of this paper. For our purposes, we chose to adopt a pragmatic approach that is based on what is currently being marketed to consumers as ‘nonfiction VR’; accepting that this includes both 360 degree video and CGI-based experiences, and that differences in subjective judgement as to what is nonfiction are unavoidable.

Sampling Methodology

Our sample was obtained from the *Mediography of Nonfiction Virtual Reality* [4] [1]. Developed by the authors, the Mediography is a publicly available corpus of over 350 individual VRNF titles, covering the period 2012-2018. Since its launch in January 2018, the Mediography has been populated through the systematic review of leading VR content channels and the programmes of international film and documentary festivals (e.g. Sundance). We supplemented this with organic keyword search on the wider WWW via search engines and social media.

VRNF media is currently being distributed (typically free of charge) through an evolving range of platform and non-platform specific channels. These include general media aggregation websites such as *YouTube* and the *Oculus Store*, and VRNF specific portals such as *Jaunt* and *Within*. Other titles are released as standalone applications that must be downloaded and installed individually. For inclusion in the Mediography corpus (and thus our sample), titles must:

- (1) Be presented in the English language.
- (2) Include a title card, production credits or both.
- (3) Self-present as a work of nonfiction, documentary or journalism; this can be through textual description and/or inclusion in a nonfiction content programme (e.g. a documentary film festival).
- (4) Present as VR media; this can be through textual description and/or inclusion on a VR-specific content channel. This criterion is particularly relevant for 360 degree video content, recognising that such content can be viewed without a VRHMD and is not necessarily produced with VRHMD viewing in mind.

Following a six month period of search and database population, 150 titles from the Mediography were selected for detailed review. This figure constituted approximately 23 hours of content and a little over 50% of the total corpus at time of selection.

Care was taken in the selection process to ensure that our sample was broadly representative of the full corpus. Specifically, our sample included titles with release dates distributed across the period 2012-2018, with proportions in each year matching closely the overall numbers in the full corpus for that year (i.e. if 25% of the full corpus was released in 2015, 25% of our sample would be selected from 2015). The average duration of the titles in our sample (8.6 minutes) was also representative of the whole corpus (8.7 minutes).

In terms of their general composition, 83% of the titles in our sample were platform independent 360 degree video pieces, 17% were platform-dependent discrete applications and 11% required external sensors to track both head rotation and body position ('6DOF' pieces). 50% were known to have appeared on the programme of at least one international film festival, with 20 titles receiving at least one award. Again, these proportions are in line with the full corpus.

The thematic composition of the sample was broad, with pieces covering topics including migration, art, space exploration, war and climate change. All titles were free to download and view at time of writing, with six exceptions that were either paid for through the relevant app store, or sourced directly from the content owner.

4 METHOD

Design

The objectives of our analysis were: 1) to identify, collate and quantify the characteristics of VRNF content, and 2) to identify and classify the role of the viewer within a range of VRNF experiences. We describe our approach to addressing each of these objectives in turn.

Identifying the characteristic features of VRNF was achieved using a bottom-up general inductive approach (e.g. [41]). This methodology results in the generation of a coding scheme of features through continual iteration and evolution during the review process. Given the novelty of the content being investigated, and the absence of a systematic evaluation of VRNF content elsewhere in the literature, we considered this approach well-suited to our needs.

The process of developing and finalising coding labels was led by the authors, with input from an advisory group of colleagues with expertise in psychology, documentary studies, computer science and interaction design. As new phenomena were encountered, a temporary label and short definition was attached, based on what the coder felt best described what they had seen. Later in the process, these labels were reviewed through group discussion and - where possible - re-assigned with a more concise description. For example, the original label 'visual labels applied to virtual environment' was later relabelled 'diegetic visual annotation'. All 150 titles were reviewed and coded by a single researcher,

with 20% independently second-coded by a second researcher. The resulting lists were then compared, with differences resolved on a case-by-case basis through discussion.

The generation of our 10 feature categories followed a similarly iterative process and was performed after all of the individual codes had been identified and their labels finalised. To develop the final categories, a short grouping exercise was conducted with our advisory group. The final 10 categories were generated through our reaching a consensus on the smallest number of categories that were felt sufficient to describe all of our feature codes.

Classifying the role of the viewer was achieved by placing each title (post-review) within the existence / influence classification scheme proposed by Dolan and Palets [13]. The scheme classifies the role of the viewer across two dimensions, defining two levels for each. The first dimension *existence* [observant / participant] describes the degree to which the viewer is made present ("exists") as an actor within the virtual world, while the second dimension *influence* [passive / active] describes the level of control that is afforded to the viewer within the story. The characteristics of each cell in the framework are expanded upon thus:

- **Passive observant:** The viewer is essentially a ghost, playing no active role in the story. They are invisible to (and go unnoticed by) other actors in the world.
- **Active observant:** As passive observant, but with some additional means of interaction such as a gaze-based point-and-click mechanic to activate cutscenes.
- **Passive participant:** The viewer is visible to (and noticed by) other actors in the world. The viewer may be spoken to or directly addressed by other actors. The overall experience however remains passive; the viewer has no ability to affect change within the story.
- **Active participant:** As passive participant, but with the means and ability to affect change within the story and its outcome.

Apparatus and Materials

All titles were reviewed at their highest available quality using an Oculus Rift or HTC Vive VRHMD. Two headsets were required as some titles were platform exclusive. However, as the audio/visual capabilities of the two platforms are essentially identical, we do not consider there to be any meaningful difference in experience resulting from using both interchangeably.

Reviews were conducted while seated at a desk in a quiet office environment, with a swivel chair providing 360 degrees of rotation as required. Our review platform was a Dell "Alienware" laptop computer (model 15R3, Intel I7 3.8GHZ, 16GB RAM, NVidia GTX1060 dedicated graphics card). This specification meets the requirements of the 2018 NVidia VR

Ready programme, a commonly accepted standard for viewing VR media content.

Procedure

Prior to review, the most recent iteration of the coding scheme was printed as a checklist and placed on the desktop in front of the reviewer. The VRHMD was then attached and made ready for content playback. During each review, the reviewer checked off each item (tag) in the checklist as it was encountered. Any previously unseen characteristics were noted and subsequently added as new tags for future reviews. Tags were restricted to objective rather than subjective features (i.e. we did not code for narrative / visual quality etc). Where required, previously tagged titles were revisited to incorporate new tags as they came to light.

Wherever possible, completion of the coding checklist was conducted at the same time as viewing the content, either by loosening the fit of the headset to allow the checklist to be visible below the eyeline, and / or by pausing the content and temporarily removing the headset. In instances where simultaneous completion of the checklist was not possible, it was conducted immediately after viewing, with repeat viewings as required.

5 RESULTS

Our review of 150 VRNF titles resulted in the generation of a coding scheme with 64 characteristics. These features can be grouped into 10 categories: *viewer role*, *point of view*, *visual composition*, *audio composition*, *gaze manipulation*, *evidence of embodiment*, *interaction*, *locomotion*, *interpersonal space* and the *manipulation of time*. A breakdown of each characteristic and their relative frequency is shown in fig. 2.

Viewer role

In 69% of the titles reviewed, the viewer assumes a **passive observant** role for the entirety of the experience. In 10% the role is **passive participant**, 5% **active observant**, and 3% **active participant**.

We however also noted a number of cases where the viewer transitions across multiple roles within the experience. Specifically, we found examples where the viewer spends some time in a passive observant role before transitioning to a passive participant role. An example of this is the piece *The Protectors: Walk in the Rangers Shoes (2017)*. Here, the viewer spends some time observing the characters as a 'fly-on-the-wall' before later adopting the role of a character, transitioning to a first-person point-of-view.

In three titles, the viewer spends some time as a passive observant before later becoming an active observant. An example of this transition can be found in the piece *The Anne Frank VR experience (2018)*, where the viewer is presented

with an introductory video scene before being invited to explore a virtual representation of the titular character's home. Finally, we found one example where the viewer spends time as a passive observant, a passive participant and an active participant. This title, *Apollo 11 VR (2016)*, begins by watching historical video as a neutral observer, later experiencing a reconstruction of events through the eyes of a central character, before finally taking control during key events such as the landing of the lunar module.

Point of View

All of the titles were composed using one or more of three distinct perspectives or points-of-view (POV): **first person**, **fly-on-the-wall** and **omniscient**. For the avoidance of any confusion in terminology, we first define what we mean by each type of perspective in terms of VRNF:

- **First person:** The experience is viewed 'through the eyes' of person placed within the virtual world. The camera is placed at an appropriate height for the role (typically adult human, but occasionally child height or animal). Narration in a first person piece is typically diegetic - i.e. provided by another actor in the piece, or via the representation of an 'inner voice'.
- **Fly-on-the-wall:** The viewer is not embodied within the scene, and the camera is placed strategically to best capture action. This may be high above the actors, or perhaps attached to a wall or the exterior of a vehicle. The camera is invisible to the actors within the scene, creating the sense that the viewer is eavesdropping. There is no external narration in this POV: the viewer is expected to draw their own conclusions as to what is going on.
- **Omniscient:** As fly-on-the-wall, but with narration in voice-over explaining the story as it unfolds.

23% of the titles used the omniscient POV exclusively, compared to 18% that exclusively used the first person POV and 9% that were exclusively fly-on-the-wall. In the remaining 50% of titles, two or more viewpoints were used at some point during the the experience.

Visual Composition

73% of the titles were entirely or predominantly composed using **live action 360 degree video**, the majority of which was monoscopic 360 video (i.e. without stereo separation), with 25% exhibiting the more advanced **stereoscopic** form (e.g. *Space Explorers*, 2018). However, we also observed that the use of stereoscopic 360 video has grown steadily since 2015, and we anticipate this trend to continue as the availability of stereoscopic 360 cameras increases.

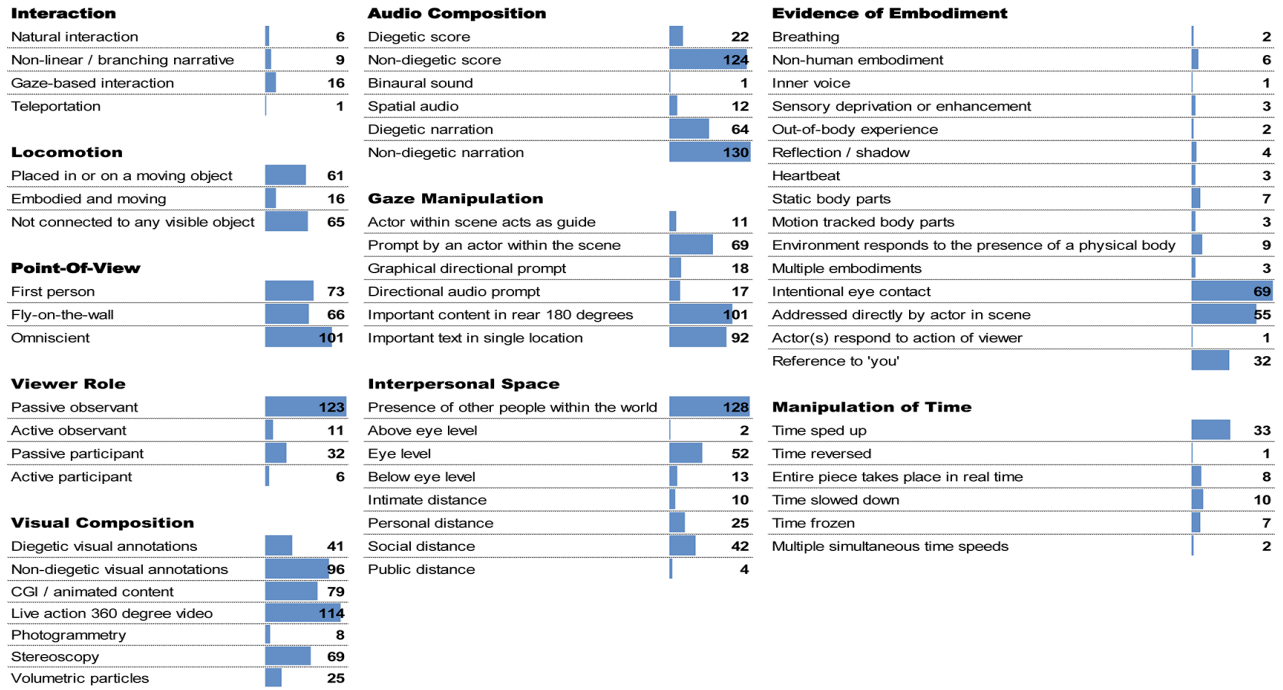


Figure 2: Characteristics of VRNF titles (n=150) by frequency of occurrence.

Just over half (53%) of the titles included at least some **CGI / animated content**, with 27% being composed entirely using animation / CGI. In titles that are entirely CGI, we observed a range of techniques that blend artistic CGI and the digital recreation of actual people and places in different ways. In some titles for example, the reconstruction of people and places are both completely computer generated (e.g. *Hunger in LA*, 2012), while in others, a real-world environment is digitally reconstructed from volumetric scanning using **photogrammetry** (e.g. *Nefertari: Journey to Eternity*, 2018). Photogrammetry is also used to digitally recreate live actors as holograms, allowing them to be placed in a computer generated environment (e.g. *Buzz Aldrin: Cycling Pathways to Mars*, 2017). A particularly striking use of this technique is the piece *After Solitary* (2017), where the digital hologram of a former prison inmate is placed within a digital reconstruction of his previous cell, allowing him to talk about his experience of solitary confinement ‘in-situ’.

To support narrative, 63% of the titles included **non-diegetic graphical annotations** (such as overlaying the name of a character who is speaking), while 27% included **diegetic graphical annotations** such as text appearing on walls (e.g. fig. 3).

Finally, we noted 25 examples of **volumetric particles**, where stereoscopically separated dust, rain or other particles

were used for artistic effect (e.g. *Notes on Blindness: Into Darkness*, 2016).

Audio Composition

The use of audio in VRNF is predominantly limited to stereo sound. Only 8% appeared to use a **spatial audio** soundfield, where sound is both omnidirectional and responsive to movement of the head. Examples with spatial sound include *Easter Rising: Voice of a Rebel* (2016) and *I am a Man* (2018). One further example, *Notes on Blindness: Into Darkness* (2016), exhibited **binaural sound**, where sound is omnidirectional but does not respond to movement of the head.

82% included a **non-diegetic score**, where background music did not emanate from within the virtual world but was instead overlaid upon it. Titles with a **diegetic score**, where the score emanates from within the virtual world, were relatively rare (14%). Following a similar pattern, 86% used **non-diegetic narration** (i.e. by voice over), while 43% used **diegetic narration** provided by an actor within the world.

Gaze Manipulation

In 66% of titles, the viewer is **encouraged to rotate their body to see content that is placed in the rear 180 degrees of view** at some point during the experience. Of these pieces, 61% included **important text that was placed for a**



Figure 3: Diegetic visual annotations. 6x9: *A Virtual Experience of Solitary Confinement*, 2016.

short time in a single location. The Oculus Rift and HTC Vive headsets used for our review both provide a horizontal visual field of 110 degrees. Though this compares well to the human binocular visual field (approx. 120 degrees), the full horizontal human visual field including the monocular / peripheral region reaches approx. 200 degrees [20]. With such a restricted level of peripheral vision, failure to manage where the viewer is looking in VRNF content can easily result in key information being missed.

We identified three general strategies that producers appear to use in order to direct the visual attention of the viewer. The first and most common is a **prompt by an actor within the scene** (46%). Typically, this is made manifest by an actor moving through the scene in such a way that the viewer feels it obvious to follow their direction of travel. In 7% of cases, an **actor within the scene takes the role of a guide**, pointing and/or using deictic language (“look at that over there”).

The second most frequently observed technique for gaze manipulation is through a **graphical directional prompt** (12%). Here, an oft-seen example is the offsetting of subtitles to sit slightly offscreen so that the viewer is gently guided in the desired direction for the next scene in order to read them. Lastly, 11% exhibited a **directional audio prompt**, such as a loud noise emanating from outside of the current frame.

Evidence of Embodiment

By “evidence of embodiment”, we here refer to mechanisms by which the viewer is provided with a physical body in the virtual world or - in the absence of a visible body - other evidence that suggests that they have a physical presence within the virtual world.

Our review revealed that to be provided with visible body parts in a VRNF piece is rare. We found only 10 titles where the viewer is provided with a body and/or limbs. In three

of these instances, **body parts (typically the hands only) were motion tracked** using the external sensors and hand controllers supplied with the VRHMD. In the remaining seven cases, visible **body parts were rendered immobile / static**.

For six of the ten cases where the viewer is visibly embodied, the body provided is clearly not intended to be that of the viewer’s own. This is either by being obviously **non-human**, or because they are experiencing the point of view of a specific character within the story. In the remaining cases, body parts are masked by a uniform such as a spacesuit.

In several cases, we observed the use of **sensory deprivation or enhancement**, where the ability to see or hear is modified in some way to support the role. Examples of this include the of visualisation of an audioscape world (*Notes on Blindness: Into Darkness*, 2016), or blurring to represent the visual impairment of the central character (*See like Menna*, 2018). In one piece the inner voice of the embodied character is heard (*The Party: A Virtual Experience of Autism*, 2017). Finally, we observed several cases where the viewer experiences **multiple embodiments** or undergoes an **“out of body” experience**, visibly floating away from their virtual body.

Beyond providing visible body parts, producers have included a range of other cues that suggest embodiment. Those include the casting of a **reflection or shadow**, **breathing**, and an audible **heartbeat**. Elsewhere, clearly **intentional eye contact** is a commonly seen external source of evidence that the viewer holds a presence within the scene. This is typically reinforced by **actors within the scene speaking to / addressing the viewer directly**, with additional evidence provided by **actors making reference to “you”** as someone they can clearly see. However, we found only one instance - *Blindfold* (2017) - in which **actors respond to actions taken by the viewer**.

Finally, we noted a very subtle indication of physical embodiment in that in several pieces, we identified moments where the **environment responds to the presence of a physical body**. An example of this is tall grass that parts as “you” move through it.

Interaction

By “interact”, we are here referring to mechanisms by which the viewer is provided with agency in the piece beyond movement of their head.

Only 15% of titles provided any additional interactivity beyond head tracking. Of these, we found that most (70%) provided all additional interactivity through a **gaze-based interaction** mechanic, where the viewer is provided with a target reticule to enable a point-and-click function. Sometimes this mechanic was used to permit the selection of



Figure 4: Branching narrative gaze selection. *Decisions: Party's Over* (2018)

objects and / or to activate a cutscene, while other examples allowed viewers to make simple binary decisions within a **non-linear / branching narrative**. An example of such selection is seen in *Decisions: Party's Over* (2018, fig. 4).

We observed six examples where the viewer was provided with motion tracked virtual limbs and / or the ability to interact with the virtual environment using **natural interaction** gestures such as the grasping, pulling and pushing of objects. In all but one of these cases, the rendering of limbs was limited to hands only (e.g. *I Am a Man*, 2018), with *Home: An Immersive Spacewalk Experience*, (2016) being the only piece we found that rendered limbs above the wrist.

Locomotion

In 29% of the titles, the position of the viewer remains entirely static for the entirety of the experience. In the overwhelming majority of cases however, the viewer experiences locomotion within the virtual environment at some point. Almost always, this movement is managed "on rails", with the viewer having no control. Despite being a commonly used mechanic for locomotion in VR games (e.g. *Robo Recall*, *Rec Room*), we found only one example where the ability to **teleport** was used in our sample (*Nefertari: Journey to Eternity*, 2018).

In the most common form of VRNF locomotion, the viewer is **not connected to any visible object** within the scene (43%). Typically these are aerial shots, where the camera is presumably attached to a drone or helicopter that is later digitally erased. Almost as frequently (40%), the viewer is **placed in or on a moving object** in the scene, such as a car, aircraft or boat. In these cases, the camera is usually positioned such that the viewer might imagine themselves as a passenger within the vehicle. Finally, and much less commonly, the movement is presented in the first-person point of view, providing the illusion that the viewer is **embodied and moving** (11%).

Interpersonal Space

85% of the titles included the **presence of other people within the world**. In instances where a title included content that is shot from the first-person POV (72 titles in all), we recorded how personal space was approached by the director. We coded for both horizontal interpersonal distance and vertical deviation from eye-level.

For our measure of horizontal personal space, we used a subjective judgement based on Hall's zones of interpersonal space [19]. The most commonly observed range for interpersonal distance with virtual actors was at the **social distance** of around 1 metre (27%), representing a close friend. **Personal distance** of around arms length (indicative of an acquaintance) was observed in 16% of titles; **intimate distance** (the peripersonal space or whisper distance) occurred in 11%, and **public distance** (approx. 5 metres) was observed in 3% of titles.

In terms of relative eye-level, 78% of the interpersonal interactions occurred at **eye level** with the actor in the scene. In 18% of cases, the viewer is placed **below the eye level** of the virtual actor, and in two cases, the viewer is **above the eye level** of the virtual actor.

Manipulation of Time

Only 5% of VRNF titles take place **entirely in real time**. The typical VRNF experience is composed of a number of short edited scenes, often switching between multiple locations and times using simple fade / dissolve transitions. In 35% of the titles the flow of time was manipulated in some way during the experience. By far the most common was the **speeding up of time** (22%), typically through time-lapse photography. In other titles, the viewer experiences time being **slowed down, frozen** or even **reversed**.

Finally, we observed instances where the viewer is able to experience **multiple speeds of time simultaneously**, based on where they choose to look within a scene. In the piece *Isle of Dogs: Behind the Scenes in Virtual Reality* (2018), the viewer is able to experience the process of stop motion animation both from the perspective of the animators (time-lapsed) or of the characters being animated (real time).

6 DISCUSSION

In this paper, we present an analysis of a relatively new form of immersive interactive media experience *virtual reality nonfiction* (VRNF). A large representative sample of VRNF content was submitted to an inductive analysis, revealing 64 key characteristics and their prevalence within works released over a six year period (2012-2018). Simultaneously, we classified each title in terms of the role that was assigned to the viewer (observant or participant) and the degree to which they were afforded influence within the virtual world

(passive or active). This process provided us with a range of insights into the way producers of VRNF are exploring and exploiting the immersive affordances of virtual reality to create new types of nonfiction media experience. During the process of conducting our analysis, we experienced how blind people can 'see' the world through sound and what it might be like to stand on the surface of Pluto. We took part in a reconstruction of the Apollo 11 moon landing, and spent a day in the life of a Maasai warrior. VRNF is not quite a film, nor is it really theatre or a form of interactive gaming, yet it clearly shares characteristics of all three.

To return to a key motivation for this work, producers of VRNF are an interdisciplinary mix of stakeholders working with a medium where the hardware and tooling is rapidly evolving and there are as yet few best practice guides from which to work. Furthermore, as there is no 'bird's eye' view of the medium, it is difficult for producers and researchers of VRNF to understand where the medium has come from and where it is heading. Many of the characteristics we have uncovered provide evidence that VRNF producers are testing the limits of this new medium in increasingly interesting ways. For example, we found pieces that experiment with a number of sensory phenomena, including the perception of time, embodiment and sensory substitution. However - and we feel this is of direct interest to the HCI community - as often as the limits of VR are being pushed back, new challenges are being revealed. The restricted peripheral vision offered by current VRHMD for example means that it is all too easy to miss important content by simply looking the wrong way at a key moment. The potential value of immersive sound is also underexplored.

Much further research is required to understand how best to present stories in VR, particularly with respect to how audiences are approaching and reacting to this new type of storytelling. We anticipate that our Mediography research tool will help support this as we expand its metadata with the results of this investigation. For the remainder of our discussion, we delve more deeply into the current state of the art in two critical aspects of VRNF as uncovered by our research: *viewer role* and *embodiment and social interaction*.

Viewer Role

To return to the title of this paper, an early promise of VRNF is its potential as the "ultimate empathy machine": immersive experiences that - through providing viewers with the means of 'standing in their shoes' of the filmmaker's subjects - could allow audiences to experience and connect to 'the real' in ways beyond what could be achieved with traditional two dimensional film. In this vision, audiences would no longer be limited to being passive observers of the story; they could be embodied and made present within it, potentially taking a much more active, interactive role.

Perhaps the greatest surprise of our exploration of VRNF to date therefore is that the amount of content that attempts to directly fulfill this ambition is actually very small. Less than one in five of the titles we reviewed were presented entirely in the first-person perspective - surely a minimum requirement for sustaining the illusion of a physical presence within a virtual world. Although there is evidence that producers are experimenting with providing viewers with multiple roles to play within VRNF experiences, our results indicate that in general, the passive viewer role still dwarfs the active viewer role by a considerable margin, and the observant role dwarfs the participant role.

A great deal of VRNF is composed of live action 360 video, a predominantly passive medium. We might therefore expect passive viewer roles to dominate, at least in the short term. However, the degree to which the passive observant role dominated over the passive participant role was unexpected. At the macro level, our analysis seems to indicate that there is a reticence among VRNF content producers to take the leap of inviting the viewer to take a participatory role in their experiences. On the contrary, there is a tendency to rely on the omniscient viewpoint, where the story is simply explained to the viewer who looks on passively from afar. Further evidence of this is found in the high proportion of pieces that provide non-diegetic scores and non-diegetic "voice over" narration.

This might suggest that VRNF filmmakers currently lack full confidence to move beyond traditional storytelling structures, audience roles and filmmaking practices. Equally it might be the case that the tooling and hardware upon which VRNF is produced and consumed are currently insufficiently mature. Either way, we argue that this must be addressed if the full potential of VRNF is to be realised. There is the risk that failure to do so may inadvertently result in experiences that are actually less impactful than they would have been if presented as a non-VR experience.

The medium however remains very young, and ultimately it is our general impression that it is the youth of the medium that is the key factor. After all, the passive observant role is the role we are all used to from consuming traditional film, and we must recognise that 360 degree filmmaking (be that live action video or CGI) is both a substantial divergence from what filmmakers are used to working with, and what audiences are used to consuming. Both will need time to adjust. It is too early to detect a trend towards or against the use of the passive observant viewer role in VRNF. However we do suggest that a sustained movement away from it over the next few years would be a useful marker of a growing confidence of creatives working within this medium.

Embodiment and Social Interaction

Following our observation that only the minority of pieces utilise a first person POV, we were also surprised to find that the pieces that do take this perspective rarely provide visible evidence of physical embodiment (e.g limbs). Motion tracked virtual limbs are rarely seen, and are typically limited to hands that float free of the body. However, when motion tracked embodiment is deployed, we found that for the most part the presence of limbs adds to rather than detracts from the immersive qualities of the piece. In contrast, pieces that included visible body parts that failed to track our movement felt and looked unnatural - an effect reported elsewhere by [15].

The degree of interpersonal space between the viewer and actors within the scene generally felt appropriate: we noted for example that to set people apart at distances below the personal distance zone (the distance that would represent a good friend) was avoided, with producers of VRNF tending to favour an interpersonal distance approximating the 'social' zone (indicating a relationship at the level of an acquaintance). This was both comfortable and 'felt right'.

However, there does seem to be some potential confusion in VRNF content that is shot in an objective viewpoint that VRNF producers should be mindful of. We found multiple instances where eye contact is made and the viewer is addressed directly when the POV was clearly objective. This to us felt akin to 'breaking the fourth wall', as opposed to feeling that the actor is talking to you as a co-present participant in a shared world.

7 CONCLUSION

We have documented the emergence and early development of a new form of nonfiction media - VRNF. Through a detailed analysis, we have identified elements of an associated narrative language; one that is evolving through cautious progress. We have described how framing the viewer's role within this narrative is a key issue for interactive VRNF and we have reflected upon the importance of consistency when it comes to evidencing embodiment within a story. The wider challenge, and an exciting area for future HCI work, is how interaction design can unite nonfiction storytelling techniques with the experiential qualities of VR. The rapid evolution from prototype to industry is evidence of VR's potential as a platform for nonfiction but we argue that, as it develops, we must strive to understand its inner workings. Our work illustrates that there is much to learn from looking 'behind the curtain' but future work should also seek perspectives from audiences of VRNF and look at the contextual issues facing creators of VRNF on-the-ground.

8 LIMITATIONS AND FURTHER WORK

There are several limitations of our work that will form the focus of future research.

VRNF is an emergent format: although VRNF has been around since 2012, significant growth in content did not occur until early 2015 with the release of Samsung's *GearVR* platform. The 2018-2020 period will be crucial to developing our understanding of the potential audience for VRNF as VR adoption rises. In future work, we will continue to identify and record key VRNF content in the VR Mediography resource, reflecting upon and evolving our inclusion criteria for nonfiction VR as the medium continues to expand.

Platform limitations: all of our reviews were conducted at the highest possible quality using the current (mid 2018) top tier of VR hardware. We are mindful that the proportion of the potential audience for VRNF that have access to this equipment is small, relative to mid-tier VR equipment such as the Google *Daydream* that offer smaller fields of view and less processing firepower. In future work, we will expand our review platforms to take account of potential differences in experience that result from using VR platforms with different audio/visual capabilities.

Audience involvement: though we have strived to maintain an objective viewpoint (while occasionally voicing our subjective suspicions), we do not know at this point how audiences approach and interpret alternative viewer roles. In future work, we will be working with audiences to compare and contrast key pieces of VRNF that have been selected based on their characteristics as uncovered in this work.

9 ACKNOWLEDGEMENTS

This research is funded by the UK EPSRC as part of the research project Virtual Realities - Immersive Documentary Encounters (ref: EP/P025595/1). For more information on the project and its partners, visit our website at <http://vrdocumentaryencounters.co.uk>.

REFERENCES

- [1] 2018. Non-Fiction VR: A Mediography. <http://vrdocumentaryencounters.co.uk/vrmediography/>. (Accessed on 09/17/2018).
- [2] Judith Aston, Sandra Gaudenzi, and Mandy Rose. 2017. *i-Docs: the evolving practices of interactive documentary*. Columbia University Press.
- [3] Jeremy Bailenson. 2018. *Experience on Demand: What Virtual Reality Is, how it Works, and what it Can Do*. WW Norton & Company.
- [4] Chris Bevan and David Green. 2018. A Mediography of Virtual Reality Non-Fiction: Insights and Future Directions. In *Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 161-166.
- [5] Paul Bloom. 2017. *Against empathy: The case for rational compassion*. Random House.
- [6] Jessica Brillhart. 2016. The Language of VR. <https://medium.com/the-language-of-vr/tagged/editing>. (Accessed on 09/17/2018).

- [7] Andy Brown, Jayson Turner, Jake Patterson, Anastasia Schmitz, Mike Armstrong, and Maxine Glancy. 2017. Subtitles in 360-degree Video. In *Adjunct Publication of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video*. ACM, 3–8.
- [8] Thomas J Csordas. 1994. *Embodiment and experience: The existential ground of culture and self*. Vol. 2. Cambridge University Press.
- [9] Barney Dalgarno and Mark JW Lee. 2010. What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology* 41, 1 (2010), 10–32.
- [10] Eric Darnell and Michael Hutchinson. 2016. Invasion!: Exploring Empathy and Agency in VR. In *ACM SIGGRAPH 2016 VR Village (SIGGRAPH '16)*. ACM, New York, NY, USA, Article 10, 2 pages. <https://doi.org/10.1145/2929490.2929494>
- [11] Nonny De la Peña, Peggy Weil, Joan Llobera, Elias Giannopoulos, Ausiàs Pomés, Bernhard Spanlang, Doron Friedman, Maria V Sanchez-Vives, and Mel Slater. 2010. Immersive journalism: immersive virtual reality for the first-person experience of news. *Presence: Teleoperators and virtual environments* 19, 4 (2010), 291–301.
- [12] Tilman Dingler, Kai Kunze, and Benjamin Outram. 2018. VR Reading UIs: Assessing Text Parameters for Reading in VR. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, LBW094.
- [13] Devon Dolan and Michael Parets. 2016. Redefining the axiom of story: The VR and 360 video complex. *Tech Crunch* (2016).
- [14] Phillip Doyle. 2017. Embodied and Disembodied Voice: Characterizing Nonfiction Discourse in Cinematic-VR. In *International Conference on Interactive Digital Storytelling*. Springer, 363–366.
- [15] Harry Farmer, Lara Maister, and Manos Tsakiris. 2014. Change my body, change my mind: the effects of illusory ownership of an outgroup hand on implicit attitudes toward that outgroup. *Frontiers in psychology* 4 (2014), 1016.
- [16] Constance Fleuriot and Jonathan Dovey. 2012. *Pervasive media cookbook*. (2012).
- [17] Nick Fraser. 2012. *Why documentaries matter*.
- [18] Arnau Gifreu-Castells. 2014. Mapping trends in interactive non-fiction through the lenses of interactive documentary. In *International Conference on Interactive Digital Storytelling*. Springer, 156–163.
- [19] Edward T Hall. 1963. A system for the notation of proxemic behavior. *American anthropologist* 65, 5 (1963), 1003–1026.
- [20] David B Henson. 2000. *Visual fields*. Butterworth-Heinemann Medical.
- [21] Jonathan David Hook. 2018. Facts, Interactivity and Videotape: Exploring the Design Space of Data in Interactive Video Storytelling. In *Proceedings of the ACM TVX International Conference on Interactive Experiences for Television and Online Video*. Association for Computing Machinery (ACM).
- [22] Henry Jenkins. 2006. *Convergence culture: Where old and new media collide*. NYU press.
- [23] Scott Barry Kaufman and Carolyn Gregoire. 2016. *Wired to create: Unraveling the mysteries of the creative mind*. Penguin.
- [24] Jarrod Knibbe, Jonas Schjerlund, Mathias Petráeus, and Kasper Hornbæk. 2018. The dream is collapsing: the experience of exiting VR. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 483.
- [25] Jaron Lanier. 2017. *Dawn of the new everything: A journey through virtual reality*. Random House.
- [26] Brenda Laurel. 2016. What is virtual reality? *medium.com*. Online article accessed August 9th (2016).
- [27] Scott W McQuiggan, Jonathan P Rowe, and James C Lester. 2008. The effects of empathetic virtual characters on presence in narrative-centered learning environments. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1511–1520.
- [28] Tassos A Mikropoulos and Antonis Natsis. 2011. Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education* 56, 3 (2011), 769–780.
- [29] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.
- [30] Chris Milk. 2015. How virtual reality can create the ultimate empathy machine. *How Virtual Reality can create the ultimate empathy machine* (accessed 22.10. 2015) (2015).
- [31] Janet Horowitz Murray and Janet H Murray. 2017. *Hamlet on the holodeck: The future of narrative in cyberspace*. MIT press.
- [32] Kate Nash. 2018. Virtually real: exploring VR documentary.
- [33] Bernard R Robin. 2008. Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into practice* 47, 3 (2008), 220–228.
- [34] Katja Rogers, Giovanni Ribeiro, Rina R Wehbe, Michael Weber, and Lennart E Nacke. 2018. Vanishing Importance: Studying Immersive Effects of Game Audio Perception on Player Experiences in Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 328.
- [35] Mandy Rose. 2018. The immersive turn: hype and hope in the emergence of virtual reality as a nonfiction platform. *Studies in Documentary Film* 12, 2 (2018), 132–149.
- [36] Mandy Rose. 2018. Technologies of seeing and technologies of corporeality: Currents in nonfiction virtual reality. *World Records* 1, 1 (2018), 01–11.
- [37] Martijn J Schuemie, Peter Van Der Straaten, Merel Krijn, and Charles APG Van Der Mast. 2001. Research on presence in virtual reality: A survey. *CyberPsychology & Behavior* 4, 2 (2001), 183–201.
- [38] Mel Slater. 2003. A note on presence terminology. *Presence connect* 3, 3 (2003), 1–5.
- [39] Mel Slater and Sylvia Wilbur. 1997. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 6, 6 (1997), 603–616.
- [40] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 289.
- [41] David R Thomas. 2006. A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation* 27, 2 (2006), 237–246.
- [42] Huyen TT Tran, Nam Pham Ngoc, Cuong T Pham, Yong Ju Jung, and Truong Cong Thang. 2017. A subjective study on QoE of 360 video for VR communication. In *Multimedia Signal Processing (MMSP), 2017 IEEE 19th International Workshop on*. IEEE, 1–6.
- [43] William Uricchio. 2018. VR: Between Hope, Hype, and Humbug. <https://lareviewofbooks.org/article/vr-between-hope-hype-and-humbug/>. (Accessed on 09/17/2018).
- [44] Angela Watercutter. 2018. A VR Movie Set in Space Just Landed a 7-Figure Deal at Sundance. Yes, You Read That Right. <https://www.wired.com/story/vr-film-spheres-huge-sundance-deal/>. (Accessed on 09/17/2018).
- [45] Brian Winston. 2002. *Media, technology and society: A history: From the telegraph to the Internet*. Routledge.