STAGE-DIRECTING THE VIRTUAL REALITY EXPERIENCE:
DEVELOPING A THEORETICAL FRAMEWORK FOR IMMERSIVE LITERACY

HENRY MELKI*
**Henry Melki** is currently the Visual Effects and Animation lecturer at Ulster University. His PhD investigated the creative processes behind believable photorealistic animated characters in film. Henry’s research interests also include areas in design theory, 3D design, Character Design, Visual Effects, Virtual Reality, Perception Theory, Empathy and Aesthetic theories.

**Corresponding Author**
Henry Melki  
[hmelki@ulster.ac.uk](mailto:hmelki@ulster.ac.uk)  
Course Director BDes Hons Animation  
Lecturer in Animation and VFX  
Ulster University  
York St,  
Belfast  
BT15 1ED

Paper submitted: 9th April 2021  
Accepted for Publication: 7th October 2021  
Published online: 17th December 2021
Abstract

Despite the incremental improvement and inclusion of immersive technologies in entertainment, training simulation, fine art, inclusive design, academia, and education; Virtual Reality (VR) still faces issues regarding its ability to compete with films and animation in visual storytelling without merging into the realm of video games.

In 2015, Pixar's Ed Catmull warned moviemakers that Virtual Reality is "not storytelling" and argued that the linear aspect of narratives poses an obstacle that cannot be overcome with VR. In contrast, Catmull argued that VR has immense application in games. However, VR creators have been pushing the boundaries and possibilities of delivering narratives in virtual spaces. In 2019, the VR experience "Gloomy Eyes" was presented at the Sundance festivals featuring a 30-minute story split between 3 episodes. The simulation is structured to provide its audience with some degrees of freedom while guiding them intuitively through the virtual space. In 2021, Blue Zoo also released a VR project titled "The Beast" featuring a cyclist powering up a snow-covered mountain. The short film was entirely created in Quill VR with the intention of being treated like a theatrical play rather than a film.

While the creators of "The Beast" have explicitly mentioned the influence of theatre, "Gloomy Eyes" draws its visual language from similar theatrical roots. This paper argues that VR has been mistakenly compared to film and animation when it should be associated with theatre. The audience of both are not passive as they are during the screening of a film or animation. The space and the medium demands participation through their presence in the same space with the actors/characters. Theatre presents a promising candidate for extracting criteria that could be used to develop a visual language for VR.

This research aims to formulate a framework for developing a VR visual language through comparison between character-driven narratives in VR such as "Gloomy Eyes" and "The Beast". The comparative study establishes overlapping criteria and characteristics found in the structure, literacy, sound, and delivery format of narratives in a theatrical performance. These criteria are then outlined and discussed, drawing from affordance theory and discussions on aural and visual attention in theatre, to form a holistic view in approaching VR literacy.

Keywords: Theatre, Virtual Reality, Immersion, Visual Literacy, Affordance
Introduction

It is undeniable that Virtual Reality (VR) has made considerable impact on society. It has demonstrated its versatile ability to enhance our experience and engagement with digital content in medicine, training simulations and education (Ghost Productions, 2019; Rutledge et al., 2019, pp. 2051-2059), entertainment (Lang, 2020), research and development (Clavin, 2019), as well as helping users develop a deeper understanding of their identity (Bissonnette, 2019, pp. 238-263).

However, despite the presented specialised innovations in VR, storytelling remains one of the main hurdles that developers are continuously attempting to overcome. While there have been interesting attempts at producing storytelling VR experiences, some lean closer to games, while others still require a more developed and unified visual language (Urban, 2021, pp. 107-127; Ward, 2017; AWN Staff Editor, 2020). This paper, consisting of four main sections, explores the historical development of visual languages of film, theatre, and games and aims to establish a theoretical framework that outlines the visual language suitable for a VR focused storytelling context.

The second section explores the development of visual literacy of film and theatre while highlighting their historical interconnected relationships. This section aims to highlight the key visual characteristics that help define the visual literacy of film while discussing the difficulty of their application in VR. Furthermore, this section identifies different approaches to storytelling, narratology, and ludology, relevant in VR that are rooted in film and games, to consider the development of narrative focused VR experiences and the outline of key principles.

The third section focuses on the application of the outlined key principle on two case studies of immersive VR experiences that are explicitly influenced by theatre, as well as other mediums such as film and games. The final section discusses the outcomes of the analysis and the opportunities for future development.

1. Technical and Theoretical Background

The possibility of interactive narratives has been one of the main focal points of discussions within the sphere of digital and immersive technologies (Lanier & Upadhye, 2017; Higgins, 2016, p. 88; Manovich, 2001, p. 249). The possibility of the viewer interacting with a constructed reality through a machine, or a tool has been considered and explored at the inception of Cinema (Higgins, 2016, p. 88; Rickitt, 2000, pp. 10, 24-25). However, the term Virtual Reality (VR), as it is known in its current digital medium, was coined in 1989 by Jaron Lanier. Since then, creative practitioners have continuously attempted to achieve a sense of presence within their work and artwork (Higgins, 2016, p. 88; Steuer, 1992, p. 73). This paper does not aim to outline the historical development of VR. However, it focuses on developing a theoretical and technical understanding of the medium before exploring its potential visual language.

Throughout the development of VR, the definition of the term has primarily centred around technology and hardware. Conducting a quick search online on the definition of virtual reality, results in recently written web pages that discusses VR from a device-driven standpoint (O’Boyle, 2021; Bardi, 2020; Mitchell, 2020). However, Steuer (1992, pp. 79-84) argues that a device-driven definition of VR is unacceptable when the medium is used for communication, research, software development, and media consumption.
The issue, with defining VR in a device/hardware centric format, is that it does not provide any clear outline of the creative processes and conceptual frameworks through which design and problem-solving decisions, of consistent quality can be made. Consequently, the medium fails at providing the users the ability to use their experiences of visual languages in other mediums to navigate VR effectively (Steuer, 1992, p. 73). Therefore, it is crucial to investigate the various components that could help formulate a holistic definition of VR which could help outline the different overlapping visual languages from other mediums.

The formulation of a Virtual Reality definition without reference to hardware is possible by exploring the theory behind the viewer's experience of presence within the constructed reality or “telepresence” (Steuer, 1992, p. 76). Steuer (1992, pp. 76-80) argues that a telepresence centric definition is informed by both the properties of the technology, and the perceiver's experience. Therefore, the resulting understanding of VR is more holistic and could clarify the opportunities the medium can offer. In a medical experiment, Ahmadpour et al. (2020, p. 8) devised a holistic approach to ensure an effective design in managing pain and anxiety in children using VR interventions. The VR design was informed by the continuous user interaction, the story told by the participant, the emotional impact of the VR product on the user, the goals that must be achieved, and the perceptual affordances provided by the technology.

The holistic understanding of the medium also clarifies the developer/viewer roles and expectations, facilitates the suspension of disbelief, encourages participation and immersion in the presented virtual reality (Ahmadpour et al., 2020, pp. 7-9; Tooby & Cosmides, 2001, pp. 10-13).

Furthermore, the definition of VR through telepresence is crucial in understanding the characteristics of the medium and the development of its visual language. The individual's perception of reality is highly dependent on their experience and inseparable relationship with their environment (Warren, 2020, pp. 151-173; Wages, Grünvogel, & Grützmacher, 2004, pp. 216 - 225; Gibson, 1986, pp. 7 - 15). The environment and the available objects in it often provide tacit information, regarding navigation, usability, and function, to the perceiver known as affordances (Gibson, 1986, pp. 33-44). The combination of these affordances presented within the virtual space, including multi-sensory stimuli, the type of interactive elements, and the characteristics of the participant, play a vital role in the perception and evaluation of the presented environment (Norman, 2013, pp. 1-36; Steuer, 1992, pp. 79-81; Gibson, 1986, pp. 47-63).

Therefore, this research examines Virtual Reality holistically by considering the cognitive perception and the experience of telepresence in the Virtual environment (from this point will be referred to as presence), as well as the opportunities provided by the technology and its limitations. Steuer (1992) relied on vividness and interactivity to discuss the holistic approach to defining VR. Vividness relates to the way the virtual environment presents information to the viewer. A highly vivid virtual environment extends to one or many of the viewer's senses through the simultaneous presentation of complex sensory dimensions, as well as the resolution or the quality of these dimensions. The communication to the sensory dimensions could be achieved through the five perceptual systems (Steuer, 1992, p. 81):

1) The basic orientation system (responsible for the body equilibrium)
2) The auditory system
3) The haptic or Touch system
4) The taste/smell system
5) The visual system

To achieve a high resolution and consistency of quality and ensuring immersion, researchers have been continuously experimenting and developing the hardware and technical requirements to establish a consistent practical pipeline in the development of immersive VR. For example, developers often
aim to provide real time graphics with high frame rates (Zielinski, Kopper, Rao, & Sommer, 2015, pp. 19-26). However, this is not always possible due to the technical aspect of the simulation and its rendering. In another experiment, Zielinski et al. (2015, pp. 19-26) explored potential solutions using low frame rate and low persistence displays techniques. The research concluded that the low persistence technique does not add any performance enhancements in the navigation and selection tasks in the VR simulation. However, users preferred using the strobing low persistence over the high persistence conditions, as it reduced nausea and sickness.

VR motion sickness remains one of the main issues that need to be addressed when using Head Mounted Displays (HMDs), especially that nausea is more prominent and frequent with women (Volpicelli, 2016, p. 21). In an experiment involving two groups, each consisting of 18 men and women, demonstrated that 14 out of 18 women felt sick during the VR simulation, while 6 out of 18 men experienced sickness (Munafo, Diedrick, & Stoffregen, 2017, pp. 889-901). The experiment suggests that VR motion sickness could be linked to the small movements human bodies make when standing still. Those who experienced sickness had unstable stances during the simulation. In contrast, Bas Roker et al. (2021, pp. 1-11) disagreed, arguing that motion sickness is due to whenever the eyes and the balance system present conflicting information. The clashing of information varies in severity between individuals, especially with those who possess greater sensitivity to varying visual cues as they are more likely to detect conflict. There was little evidence supporting the differences between Male and Female users. However, the researchers hypothesise that the reported the motion sickness differences between Men and Women is due to poor personalisation of VR displays. This makes the experience exclusive, which hinders full immersion for some users. The identification of the stimuli that trigger reactions from participants with greater sensitivity is key in developing personalised VR experiences and mitigate discomfort (Fulvio, Ji, & Rokers, 2021, pp. 1-11). A key stimulus that Bas Roker et al. (2021, pp. 1-11) have identified, is the sensitivity to motion parallax which could result in motion sickness. This phenomenon occurs when moving objects travelling at a constant speed appear to move greater amount if they are placed closer to the viewer or the camera than those placed at a further distance (Brinkmann, 2008, p. 46).

In addition, other technological and user experience enhancements in VR also present an important element in maintaining engagement and immersion. These enhancements include, for example, the development of the communication systems within VR to ensure high quality untethered VR streaming (Chakareski, 2020, pp. 6330-6342).

These developments are very important to the continuous evolution of VR and the quality of its delivered simulations. The consideration of motion parallax is a vital parameter to consider when designing a personalised display in VR. To ensure that the visual literacy of VR is developed through a holistic consideration of parameters, it is also important to investigate the visual conceptual side of VR to explore the potential development of its visual language for narrative driven virtual environments.

2. Visual Literacy of Film and Theatre

The incremental technological improvement has, without a doubt, enhanced the viewing and interactive immersive experience of VR. However, Virtual Reality does not have a standardised visual literacy for visual storytelling that allows it to compete with films and animation. A unanimous agreement on the storytelling principles of the medium, and whether they could be achieved consistently in VR, require further development. Pixar’s Ed Catmull, for example, argued that while Virtual Reality has great potential in games, it is “not storytelling” because the linear aspect of narratives poses an obstacle that cannot be overcome with VR (Dredge, 2015). Similarly, debates on whether games are a narrative based medium, have been discussed where the meaning of narratives have shifted depending on the use of narratology, defined by the
sequence and rhythm of events, or ludology, where the player has a defined degree of control over the progression of the story (Aarseth & E., 2012, pp. 129-133).

Case studies present a considerable number of attempts at developing storytelling-based simulation in VR fall under ludology approaches and lean closer to games or training simulations that use storytelling a secondary element to enhance learning outcomes (Urban, 2021, pp. 107-127; Zhang & Bowman, 2021, pp. 326-329; Ulaş, 2014, p. 2040). The use of ephemera as a tool for storytelling, as mentioned in Urban’s (2021, pp. 107-127) research for example, still relies on the player’s interaction for the story to progress which would comply with ludology principles rather than narratology.

2.1 Visual Literacy of Film

To establish the principles or a framework for a visual language for VR immersive experiences that are purely narrative based, this paper examines the development of the visual literacy of other storytelling mediums. This section considers film, whether live action, animation, digital or traditional, as an important medium to explore in the development of narrative focused VR experiences.

Cinema has been at the forefront of visual entertainment and storytelling for generations. Its roots can be traced to different origins including magic lantern shows, theatre, shadow performances, and animation (Manovich, 2001, p. 82; Rickitt, 2000, pp. 8-13). Initially, these mediums provided limited content that struggled at exciting the audience’s imagination. Therefore, audiences preferred theatre over hand-drawn projections such as magic lantern shows (Rickitt, 2000, p. 8). However, with the innovations that occurred in the 19th century, especially with the developments in photography which allowed the recording of live-actor performances on film, technology became one of the main driving forces that continuously altered the perception and understanding of film (Manovich, 2001, pp. 88-89).

Previously, film and animation were two separate mediums. Film focused primarily on the use of photographic film to record reality, a representation of reality, and live actors/individuals. In contrast, while the principle of presenting the sequence of frames is similar to film, animation focused on hand drawn illustrations presented in sequences to create an illusion of life and movement (Webster, 2002, p. 1; Wells, 1998, p. 10). As technology developed, the advent of digital cinema resulted in the merger of film and animation.

Currently, most films consist primarily of digital and composit footage that is generated and animated using computers (Al-Jamea & Rizvi, 2017; Leberecht, 2014). The merging of media resulted in unforeseeable issues regarding the mixed use of computer-generated imagery (CGI) and practical effects, and the increasingly competitive and fast-paced industry, that required the formulation of standardised creative process for an effective and sustainable delivery of projects (Al-Jamea & Rizvi, 2017; Leberecht, 2014).

Similarly, VR, as a relatively new medium, has not achieved a defined standardised framework that can be used to deliver VR narratives without turning the simulation into a game (Fearghail, Knorr, & Smolic, 2019, pp. 1-8; Damiani & Southard, 2017). As with any visual medium, the main issue that needs to be addressed is the method or the design process through which the viewer is able to read the visuals as the creator intended (Kress & Van Leeuwen, 2006, pp. 4-6).

In film, the director can use depth of field, lighting, framing, colours, compositions, and other camera and cinematography techniques to lead the attention of the viewer to the intended area (or salient elements) in the frame to effectively communicate the narrative (Fearghail et al., 2019, p. 1). The challenge with VR, in contrast to film, is the degrees of freedom possessed by the audience. When viewing a film, the audience, while engaging with the presented content, remain passive and is not provided any affordances through which they are required to interact with the film.
To read or analyse a film, visual methodologies need to be considered to outline the list of criteria that breakdown the visual language’s main components in the shots (Rose, 2016). Throughout the book *How to Read Film: Movies, Media, and Beyond*, James Monaco (2009) outlined the list of parameters needed to analyse the various shots of a film. These parameters are summarised and listed under three categories below:

The first category is the *mise-en-scène* which can be discussed through two main elements, the frame, and the shots. The components of these two elements are outlined below (Monaco & Lindroth, 2009):

**The Frame:**
- screen ratio (for example, the ratio of 16:9 is used for HDTV).
- open or closed screen framing.
- the screen plane which includes the frame plane, geographical plane, and depth plane.
- Multiple images.
- Superimpositions.

**The Shots:**
- Shot distance which includes close-up shots, head and shoulders shots, and long shots.
- Shot focus.
- Shot angle.
- Point of View.
- Pans, Tilts, Rolls, and Zooms.
- Tracking and Crane shots.

The second category focuses on the montage of the moving image, which is outlined with the cuts, their types (for example, fade in/out, unmarked cuts, jumps, dissolve, and iris), and their rhythm.

The third category, Monaco (2009) focused on the use of sound regarding its type (for example, music, speech, environmental sounds) and its relation to the image (for example, the sound’s source, whether it is parallel or contrapuntal).

While Monaco’s defined parameters were defined in the context of film, they also apply to animated films. Since animated characters and avatars play a crucial role in VR as well, the principles of character design and animation need to be considered in establishing an understanding in presenting animations in the context of VR. For example, the 12 principles of animation are crucial elements in visual storytelling that help the animator convey a believable character performance (Thomas & Johnston, 1995, pp. 47-70). Furthermore, the following elements are vital in the achieving saliency on screen and communicating the context and the story of the characters on screen (Publishing & 3dtotal, 2020, pp 106-119; Tillman, 2011, pp. 11-23, 67-84; Kress & Van Leeuwen, 2006):
- Archetypes
- Silhouettes
- Shapes
- Poses
- Colour
- Hierarchy
- Orientation
- Composition

In western cultures, for instance, movements (character or camera) occurring from left to right seem more fluid and natural, while those moving from right to left are perceived as unnatural or tense or stressful (Kress & Van Leeuwen, 2006, pp. 4-6). In addition, a character’s silhouette and proportions could indicate to the audience the expected performance and personality, whether it is the sinister appearance of a villain, or the benevolent and empowered stance of the protagonist (Publishing & 3dtotal, 2020, 236-241; Tillman, 2011, pp. 67-84).

In VR, even if the content does not require interaction, the availability of affordances and the degrees of freedom intuitively
lead the viewer to become an active participant (Steuer, 1992, p. 84). While the environment could be forced to be confined within a frame, VR has unique characteristics that should be explored rather than imitating existing mediums like cinema. Therefore, some of Monaco’s criteria, the screen ratio, cuts, and camera angles for example, should be reconsidered and replaced with other criteria that are more relevant to VR. Therefore, the visual literacy parameters could be extracted from Monaco’s criteria and adjusted to meet the requirements of the medium.

Due to the active nature of the viewer in VR, certain degrees of freedom must be considered. The Virtual environment no longer exists within the frame of a screen. This allows the viewer to move and perceive the content from different directions and angles. These degrees of freedom make the issue of saliency, or the intended viewing area where the narrative occurs, more difficult to solve (Fearghail et al., 2019, p. 1). In addition, the conservation of the holistic examination of VR literacy also requires the external, or real, environment as an added parameter that must be considered.

Therefore, the degrees of freedom must be constrained within specific parameters and design decisions that would intuitively guide the viewer through the narrative by relying on affordances as well as a variety of stimuli. While film might not entirely present the appropriate parameters for developing VR’s visual language, theatre, on the other hand, presents the viewer with enough degrees of freedom to ensure engagement, interactivity, and immersion without crossing the narrative into the category of ludology and video games. The next section discusses the various aspects of theatre and their possible application in VR.

2.2. The Visual Literacy of Theatre & The Phenomenology of Attending

Theatre’s impact on film and visual storytelling is undeniable. Whenever discussing film and its relation to theatre, it is not surprising to see the name George Méliès appear in discussions (Sontag, 1966). The association of Méliès’s films to theatre is primarily due to the reliance on the staged scenes, props and costumes, to convey imaginary journeys and objects, as well as physical metamorphosis (Sontag, 1966, p. 25).

Furthermore, theatre played a crucial role in film, years after Méliès. When sound was introduced to film and became increasingly popular, filmmakers were forced to shoot in studio closed sets, instead of the outdoors, due to the primitive microphones and sound engineering techniques (Rickitt, 2000, pp. 20-23). Unlike contemporary microphones, in the 1920s & 1930s it was difficult to isolate, edit, or remove background noise coming from street traffic or natural elements. Sets had to be built indoors with the recording equipment hidden inside or behind props on set. Filmmakers were then more inclined to hire theatre actors for their films due to their experience and ability to perform and articulate clearly and loudly on set for the sound recording (Rickitt, 2000, pp. 20-23). These occurrences had a fundamental impact on the pipelines used in film as well as their visual language.

Despite of the overlap between the two mediums, a film by Méliès, for example, cannot be considered a theatrical performance. The placement of the camera in front of the stage to record the action, the presentation and consistency of the characters through time and space are characteristically cinematic (Sontag, 1966, p. 25). Therefore, a vital difference between the two media is that once the actor’s performance is recorded and projected on screen, it features a constant presentation with each screening as opposed to the unique delivered experience in a theatrical performance (Sontag, 1966, pp. 29-31).

Furthermore, another crucial distinction between film and theatre is the use of space. Theatre features a continuous use of space. In this case, the actors are visible and in contiguity with each other on the stage space (Sontag, 1966, p. 29). In film, the use of space is discontinuous and presents a form of syntax achieved through edited shots and sequences. The
cuts, sequences, and rhythm of the shots, as Monaco (2009) also discussed, is a characteristic and important component in film (Sontag, 1966, pp. 29-30). The performance of the actors is usually defined by the edited sequence and does not exist outside the realm defined by the frame. Despite of their theatrical influence, this applies to Méliès’ films, mainly due to the performance being recorded on film and defined through the limit of the camera frame (Sontag, 1966, p. 25).

Theatre also presents a performance which has the potential to offer multiple interpretation of a mediated play. The audience is then perceiving the performance with their own eyes instead of what the camera has previously recorded (Sontag, 1966, p. 30). This means that a theatrical performance, which is intended to be experienced live, can be experienced differently with each viewing and loses its impact when recorded. Therefore, the act of attending theatre has a phenomenological implication where the presence of the audience in the theatre space plays a crucial role in the theatrical experience (Home-Cook, 2015, pp. 1-2). This is an important aspect to consider in the development of a narrative focused VR literacy principles because, like theatre, VR experiences are highly dependent on the holistic view of the presented environment with its different stimuli and affordances. The VR experience also allows the viewer to experience the simulation with slight differences each time. This also implies that, unlike film and animations, the characters cannot be shown in specific sections clipped by the camera frame. Like theatre, the animated VR characters are presented as a whole, and it is the role of the animator to ensure that the performance of the entire character comes through accurately in the simulation. For example, in animated films, the director can use perspective to create dramatic shots of the character (Publishing & 3dtotal, 2020, pp. 14-15). However, this is difficult to achieve in theatre and VR where the drama is highly dependent on the characters’ performance, music, sound, light, colours, and contrast.

As Steuer (1992, pp. 80-81) mentioned, to achieve a highly immersive experience in VR, the virtual environments need to extend to one or multiple of the viewer’s sensory dimensions. Therefore, the perceptual systems (see section 1) for communicating to a viewer’s sensory dimensions need to be addressed in the context of theatre.

George Home-Cook (Home-Cook, 2015, pp. 1-2) argued that “attending theatre” does not only imply being physically present in the space of the performance. There is a collective, as well as the individual, with the sense of responsibility and commitment generated by the act of attending theatre. Theatregoers are then required to make an effort and participate by "stretching" themselves to the performance. As a response, theatregoers make themselves heard, they announce or draw attention to their attendance by following (or not following in some cases) certain protocols. For example, offering applause, producing the sound of chattering, gradually silencing their sounds as the lights dim before the start of the performance (Home-Cook, 2015, pp. 1-2). The presence of the observer in the same space as the actors/characters and actively following their performance is crucial for the phenomenological experience of attending theatre.

Therefore, unlike film, theatre is not an event that only requires attendance, but also it demands participation and attention. Home-Cook (2015, p. 3) emphasises on the importance of sound (or silence) in a theatrical performance. However, there are other stimuli that require attention that would fall within the outlined sensory dimensions in section 1.

Gerald (2007, p. 132) argued that attending theatre involves paying attention to objects, the bodies, gestures, voices, words, sounds and music that are not only present on stage in front of the audience, but also engulfed in the space around them. It is then the role of the stage director to build a live performance (Niculescu, 2019, p. 96) and plan, design, and choreograph the various affordances that would intuitively draw the attention of the audience to the salient areas in the performance, and ask them to participate (Home-Cook, 2015, p. 6).
When developing a set of standardised principles for producing narrative focused VR experiences without crossing into the category of games and ludology, theatre presents interesting insight on the criteria that can be extracted. Following the literature covered in this section, a summary or the main characteristic of theatre is outlined below.

- Theatre presents a continuous use of the space.
- The theatre has the potential in presenting the multiple interpretation of the mediated play.
- Instead of experiencing a recorded performance through a camera's lens, the audience perceive the performance with their own eyes.
- Theatre can communicate to multiple sensory dimensions through diverse perceptual systems.
- The attendance of theatre includes
  - presence in a space.
  - being part of a collective.
  - being an individual.
  - Having a sense of commitment or responsibility to the play
- Theatre demands effort from the theatregoer to pay attention to the affordances presented through:
  - Objects/props
  - The space itself. This includes the stage, the lights, the set, and other stimuli emerging from the surrounding space.
  - The music
  - The sounds & silence
  - The voices
  - The words
  - The bodies & the gestures (of the performers)
- In an animated VR context, this can be seen through the poses, posture, the silhouettes, shapes, and colours used to achieve bring the focus to the characters.
- The animation, in this case, is the transcribed performance of the animator unto the character in the chosen medium (Crafton, 2013, p. 16).

The following section aims to apply these principles, while considering the technical aspects of VR outlined in section 1, on two case studies focusing on narrative storytelling in VR. By applying these principles, the case study analysis aims to highlight the effectiveness of using theatre literacy as the basis for the VR framework for creating narrative focused simulations.


This section aims to deconstruct and analyse *Gloomy Eyes* (2019) and *The Beast* (2020), two VR case studies presenting a narrative focused approach to VR storytelling. This section places the case studies under the scope of the theatrical characteristics outlined in the previous section, while considering the main advancements and techniques required to create optimised and personalised simulations. The two case studies were chosen due to their explicit influence from diverse media including film, cinema, theatre, and games, as well as their critical reception and impact (Berry & Winn, 2021; Baker, 2020; AWN Staff Editor, 2020; Cayrol & Heller, 2019). The objective of the analysis is to pinpoint the differences and the overlapping areas of the two projects with theatre and film to formulate the main visual language that combines both.

3.1. Gloomy Eyes

In 2019, at the ACM SIGGRAPH 19, producers Antoine Cayrol and German Heller (2019) presented the Tim Burtonesque VR experience titled “Gloomy Eyes”. The VR storytelling experience presented a 30-minute story split between 3 episodes featuring influences from cinema, theatre, and spatial sound, but does not require any action from the viewer except to participate in the performance (Cayrol & Heller, 2019). The simulation is structured to provide its audience with some degrees of freedom, mainly rotating 360 degrees in place, while guiding them intuitively through the virtual space. The three-episode structure functions as the three acts of the narrative and
provides the developers room for keeping each episode/act optimised to ensure smooth and vivid display. This analysis considers that the technical considerations to achieve optimal performance for a VR experience have been achieved and focuses on the visual elements present in the VR experience.

As Home-Cook (2015) discussed, the importance of sound and silence, as a medium through which theatregoers engage and interact with the theatrical play, is crucial in theatre and is apparent with the almost synchronised diminishing lights and audience's sounds as they await the curtains to open. Similarly, as soon as the first episode is chosen, the viewer is engulfed in darkness and silence, just before the narrator begins to speak.

Another element that is highly influenced by theatre, is the use of light to draw the attention of the viewer to specific areas in the simulation. For example, once the simulation begins, the narrator begins speaking in the void, then the viewer is presented with a glowing whisp, whose contrast with the dark efficiently draws the attention to its movement which comes to a halt in front of Nena, one of the main characters. The characters’ glowing eyes, specifically the big and yellowy eyes of Gloomy and Nena, also play a crucial role in the saliency of the scene.

The use of light’s contrast, against the dark background, is consistent throughout the simulation across the three episodes. This use of contrast, shadows, and lamps does not only help drive the narrative forward, but also the transition between areas in the environment and ensure effective saliency is achieved in each scene. The viewer can always find their way navigating the environments without getting lost if their attention was drawn to other areas in the environment. The transition between locations is not achieved through cuts, but instead the setup of the environment is continuous throughout the simulation. The participants only need to rotate as the environment unfolds before them like a popup book and a play’s stage. The sound, props, objects, backdrops, and characters play a crucial role in allowing the participant to understand the context, their location, as well as the direction in which they are required to rotate. In case the viewer loses track or gets distracted looking at a specific area, the sound, which makes good use of the left and right speakers can easily guide the participant into looking towards the right direction.

In addition, the characters are also continuously performing and jumping between scenes, just as theatre performers move behind the stage to reappear once the new stage is set. The characters’ poses and body language is also reminiscent of the theatre performers, where the whole body is seen and is often facing the audience, demanding attention. In a scene explaining the Nena’s Uncle goal to hunt down zombies, one of his henchmen aims his gun at the audience while he’s looking for the undead. While logically the pose and the animation does not make sense, acknowledging the presence of an audience thematically works in achieving drama and theatre-like performance.

The gestures, poses, and performances are also accentuated by sound which is designed to attract attention as much as possible. For example, whenever the Nena’s uncle is introduced, the character is presented as tall, upper-class, serious character with glowing red eyes, a commanding presence that is difficult to miss, but also accentuated by an operatic style of off-key singing. While the character does not make eye contact with the viewer, there is still an acknowledgement of an audience. His pose continuously faces the audience throughout his performance, and his off-key singing are glaringly asking for attention.

Furthermore, in the scene where Nena meets with another crucial character in the story, labled as The Sun, the character’s hand is the only visible part that points upwards, explicitly telling the viewer to look up where Gloomy is awaiting his turn to be introduced unto the stage.
In the final part of the first episode, the Sun also engages with the viewer. The Sun acknowledges the presence of Gloomy as well as the viewer with whom the character makes eye contact and shines its light in a beam towards the viewer. The acknowledgement of the viewer is an indication of the application of theatre literacy in VR. Although the simulation only involves one individual in the viewing space, the participant is part of a collective that include the digital characters, the sound, the environment, and the props.

While the performance is recorded, the viewer still retains a degree of control regarding the areas they wish to focus on. In addition, even though the tracking system allows the environment to load at the appropriate location, based on the headset, for the viewer to engage, slight variations could occur given a slightly different experience each time. However, this might result in some motion sickness as some simulation might load with the environment loaded too close to the viewer, leading to emphasised motion parallax. The issue of motion sickness could be mitigated by providing the participant the option to adjust the distance between them and the staged environment at any moment to ensure a personalised experience suitable for their visual requirements.

In conclusion, Gloomy eyes does not only present a clear influence from theatre but also behaves like a theatrical play that can only exist in the realm of VR. It is developed to be as optimised as possible despite the dark and large presented 3D environment that could in other cases result with visual artefacts, noise, and simulation lagging. Revisiting the key parameters that define theatre outlined in section 2.2, Gloomy Eyes achieved the following:

- The simulation presents a continuous use of the space which unfolds and gets revealed as the narrative progressive.
- While there is no reinterpretation of the simulation at each playthrough, the degrees of freedom provide the viewer different experiences based on their decision on where to pay attention each time they load the simulation.
- The simulation is not live. The characters’ performance are already animated. However, the real-time rendering as well as the slight variation of the simulation’s position that occurs at each runtime, presents, to a degree, a unique performance each time the simulation loads.
- The simulation communicates to multiple sensory dimensions through:
  - the auditory system as sound was a crucial component in Gloomy eyes.
  - The basic orientation system since the viewer knows where they are located and how the simulation is going to unfold.
  - A rich visual system, despite some imperfections at times that Baker (2020) mentioned, that is well designed, featuring captivating and fluid animations, as well as visual cues that also help draw the viewer’s attention to a specific points.
- The sense of attendance is achieved through:
  - Achieving the sense of being present in the space. This is apparent at the beginning of the simulation when the environment becomes dark and silent before the narrative begins.
  - The viewer is part of a collective. However, in this case it is not with other viewers. Instead, the viewer is part of the simulation as the characters and the affordances demand constant attention from the viewer.
  - The viewer is also an individual in the simulation possessing the degrees of freedom to turn, follow, and focus on the elements they choose, or pay attention to the salient areas.
  - To follow and understand the narrative, the viewer must commit and make an effort to follow the affordances and pay attention to the simulation.
  - Just as in theatre the simulation demands attention from the viewer through objects/props, the environment, lights, music, the body language, gestures, sound, and voices of the characters.
3.2. The Beast

In 2020, Blue Zoo also released a 2D VR project titled “The Beast”, entirely developed in Quill VR, featuring a cyclist powering up a snow-covered mountain. The short film was created with the intention of being treated like a theatrical play rather than a film (Berry & Winn, 2021).

The studio conducted a lot of experimentation to figure out the pipeline in producing a 2D VR experience in a 3D virtual environment. The creators discussed their initial planning process to stage the action, create saliency in the different scenes, developing the environment and its content to be in proportion with the viewer to achieve spatial awareness (Berry & Winn, 2021). The goal of the project is to place the viewer in a storm occurring on a mountain as well as in the mind of the character cycling up to the peak of the mountain. To ensure immersion, the developers avoided frequent cuts and maintaining long shots. The developers also experimented with the use of sound, voice overs, and the use of space to create a sense of losing control (Berry & Winn, 2021).

The simulation, accompanied with the narrator, begins with the viewer floating in the sky, above the clouds, with a few flying ravens. The ravens start flying to a lower point, the clouds disperse slightly, and the cyclist becomes visible below. The viewer is then transported behind the character cycling up the steep mountain. While the cyclist is pushing forward, the narrator describes the growing storm and mentioning the rain which visually becomes more visible. The ground begins to crack, and the character looks behind while the viewer surpasses the cyclist. The gesture of the character, looking backwards, is a subtle indication for the viewer to also look behind and witness the appearance and growing size of “the beast” chasing after them, symbolising the challenges of pushing through the storm while climbing the colossal mountain. The acknowledgement of the viewer’s presence is an invitation for participation in the performance. The viewer is part of the character’s journey and a participant in the mountain climb.

The cyclist pushes forward ahead of the viewer again and the beast begins to emerge from the mountain in the form of aggressive spikey rocks and cracks in the ground as the camera shakes. The viewer is then thrown down into a dark void surrounded by eyes, and is directly addressed by the beast, indicating the state of mind of the character. Sound plays a crucial role in these scenes with the beast emerging from all sides, overwhelming both cyclist and viewer before they fall into the pit. The numerous eyes in the void, making it difficult for the viewer to have a salient area on which to focus, also add to the disorientation. The overwhelming use of visual elements and sounds emphasise the dramatic effective and the perception of the cyclist’s inner struggles and state of mind. The viewer is then also expected to struggle and keep watching just as the cyclist must keep climbing up the mountain. However, this also amplifies the potential risk of motion sickness due to the visual cues causing motion parallax. The repetitive formation and scattering of the rocks, on the mountain, and the eyes, in the void, at different scales and distance to the viewer is a possible factor for audiences to experience motion sickness.

The description of The Beast indicates a strong theatrical influence in the development of a narrative focused VR experience. It was developed to place the character and the viewer in the same frame of mind. Applying the theatre parameters outlined in section 2.2, The Beast follows these principles:

- **The Beast** presents, to an extent, a continuous use of space by relying on a long shots and seamless transitions. For example, the scene when the cyclist is climbing up the mountain and the beast begins to appear as spikes leading to the fall of the viewer in the dark void. However, there is some use of cuts and transitions that break the continuous use of space, such as fade in/outs and swipes, that are more characteristic in cinema rather than in theatre.

- Although the saliencies in The Beast are subtle, they provide the viewer space to interpret what was intended in the shots. In addition, the ending is vague and is left for the viewer to decide whether the cyclist reached the peak or not.
• The simulation allows the viewer to look around, however, there are clear indications that the viewer is not in control and experiencing the simulation through a camera. This is apparent whenever the spikes pop out of the ground and the camera shakes.

• The simulation communicates to multiple sensory dimensions through:
  - the auditory system is presented through the sounds, music, and the voiceover presented throughout the project. The voiceover plays the role as the cyclist as well as the beast. Furthermore, the sounds also function as the stimuli for drawing the attention of the viewer to specific areas.
  - The basic orientation system is applied through the simulation. It is somewhat disconnected at specific parts due to the cuts, however, this could be intentional due to the aim of the project to make the viewer feel like they are losing control especially when he falls into the void. The long shots provide the viewer with the opportunity to understand their role and orientation within the virtual space.
  - An effective visual system, where the cyclist’s red colour scheme contrasts effectively with the overall blue tint of the background. Furthermore, the camera shakes, the rain, and the overwhelming use of the eyes in the dark pit, provide effect visual stimuli to the viewer.

• The viewer’s attendance and sense of presence in the space is achieved through:
  - The viewer and the character are meant to share the same mindset. The developers aimed at placing the viewer in the head of the cyclist. This is clear when the viewer falls into the dark pit and is addressed by the beast who represents the inner struggle and challenges that the cyclist is facing.
  - The viewer retains their individuality as a spectator by having degrees of freedom which allow them to look and experience the simulation differently
  - The viewer is also responsible in committing and participating in the simulation by following the affordances and tacitly communicate with the characters. The outcome of the ending of the simulation is left for the viewer to interpret. The viewer is also invited to participate and engage with the content by looking back whenever the character turns, for example.

• For the narrative to communicate the story effectively, the viewer must participate in the story by following the available affordances presented through the gestures, such as the cyclist looking over his shoulder indicating that the viewer should also look behind them as well, the feeling of losing control in the void or feeling disoriented when the camera shakes, and finally by interpreting the outcome of the story.

4. Discussion of Outcomes & Conclusion

The review of Gloomy Eyes and The Beast has demonstrated how both case studies are, not only highly immersive VR experiences, but also very theatrical in their approach in presenting their content. Both case studies successfully delivered their visual storytelling by maintaining a narrative rather than a ludological approach. While narratology usually requires engagement and less interactivity from its audience, the two VR experiences involved the viewer in the experience and demanded participation and a sense of commitment and responsibility.

An important element that facilitated the involvement of the participant and ensured immersion was the continuous use of space. Although, Gloomy Eyes was more successful at applying this principle in their project, The Beast still provided long inclusive shots and broke them through cuts and transition to create a sense of losing control. The effective use of affordances to achieve saliency throughout the simulation is also a vital element that maintains immersion and draws the attention of the viewer to the appropriate storytelling area of the environment, enabling the viewer to successfully achieve a sense of presence.
However, both case studies did not entirely meet all the theatre principles outlined in section 2.2. *The Beast* did not maintain a continuous use of the environment throughout the simulation as there were cuts and transitions that disconnected the flow of the space and relied on rhythm of sequences as well as an explicit presence of cameras which are cinematic traits. Despite their justification for creating the effect of losing control, the discontinuity and the explicit presentation through a camera diminished the degrees of freedom which could have enhanced the immersion in the VR space. Furthermore, the lack of opportunities for multiple interpretations of the play, specifically with *Gloomy Eyes*, is a lacking theatrical trait that is missing in the simulations.

The issue with potential motion parallax cues must also be addressed. To maintain immersion, motion sickness must be addressed by providing options or the opportunity to customise the VR simulation for personalised user experience. Additional experimentation with scale and proportion is required to identify the visual cues that could trigger motion parallax phenomena and, by extension, motion sickness. Personalising the VR experience for different users should be considered as a necessity to ensure, not only immersion, but also inclusivity.

For VR to operate as its own original and independent storytelling medium, it cannot be completely theatrical or cinematic. Therefore, VR cannot be a substitute for film or theatre. It offers unique experiences that still required refinement through consistent experimentation. Developers should strive at developing its independent and standardised visual language while considering holistic parameters including usability, personalised experiences, technical optimisation, saliency, affordances, and conceptual and design frameworks. There is potential in using the parameters of visual literacy in theatre as a starting point in the development of narrative focused VR storytelling experiences. The criteria proposed in this paper could function as the foundation for a potential visual language tailored to the development of VR experiences based on narratology. These principles or parameters could be further developed and refined with additional experimentation and analysis to form a mature visual language for VR storytelling.

References


Blue Zoo, Oculus & Facebook (Producer), & Berry, G. and Winn, D. (Directors) (2020). *The beast*. [Video/DVD]


Life after pi. Leberecht, S. (Director). (2014).[Video/DVD]


Ulaş, E. S. (2014). Virtual environment design and storytelling in video games doi:10.1386/mvcr.4.1.75_1


