The Problems and Potentials of VR for Documentary Storytelling

Dirk Eitzen*

Franklin & Marshall College (United States)

Abstract

A primary goal of many VR documentaries is to foster empathy with others. Unfortunately, some of the very qualities that make viewers' experiences in the new medium unique and compelling, including the freedom to look around, also compromise its effectiveness as a means of genuinely understanding the experiences of others. This essay explores the question of whether the VR medium may have other compensatory advantages for documentary storytelling. It does this by considering social and psychological differences between third-person storytelling, which includes popular movies, first-person storytelling, as in many video games, and second-person storytelling, which is a common feature of documentaries. A key conclusion is that, even though VR is never likely to be a particularly effective instrument for creating empathy with others, through special forms of second-person storytelling, it does have some unique and effective means of engendering caring for others.

Keywords: VR and Nonfiction; VR and Narrative; VR and Empathy; Presence and Immersion in VR; Perspective or Point of View in VR.

∗ ✉ deitzen@fandm.edu

1   The Problem with VR

The special power of VR as a medium, as anybody who has experienced it will attest, is its capacity to give the viewer or user the sense of being right in the middle of a place: a refugee camp, the cockpit of a spacecraft, floating through the Grand Canyon, or whatever. This "feeling of being there" is frequently referred to as *immersion*. VR's sense of immersion derives mainly from the viewers' ability to look around freely the presented space. (By the way, unless I specify otherwise, I use the term VR in this essay to refer to both 360 VR, so called because it records places and events using a 360-degree panoramic camera, and interactive VR, with its computer-generated environments, most often used for gaming.)

When VR began to gain traction as a commercially and technically viable medium of entertainment (Google Cardboard was introduced in 2014; the HTC Vive, Oculus Rift, and PlayStation VR platforms in 2016), theorists and practitioners of the new medium had grand expectations that this power of immersion would enable a new form of storytelling, in which viewers did not merely observe characters and events, as in traditional cinema, but also had a sense of being directly involved with them. Documentarians were particularly excited by this prospect. For example, Chris Milk, the producer of the pioneering VR documentary *Clouds over Sidra* (2015), famously referred to 360 VR as an "empathy machine." He reasoned that when you are physically *with* somebody, in their presence, you feel a kind of connection to them, which translates into caring and compassion, and that this effect carries over into the virtual presence you feel with VR.

There are some pretty glaring holes in this logic, as empathy researcher Paul Bloom (2017) and others have observed. If you are standing next to a man who is laying bricks, your physical proximity to the bricklayer may help you grasp more viscerally the business of laying bricks, but it is no guarantee that you are going to feel connected to the man, much less care about him. Standing next to him virtually, in VR, isn't going to change that. And what if, instead of laying bricks, the person you are virtually next to in VR is waving a Nazi banner or having a colonoscopy? There would no doubt be some emotional impact in the "feeling of being there" in such situations, but it would be quite different from empathy.

Anyway, in spite of the hoopla surrounding VR in its early years, when it comes to documentary storytelling and storytelling generally, the new medium appears to be turning out to be something of a flop. When you look at the current offerings of With.in, which bills itself as "the premier destination for cinematic virtual reality," you find short-form experiences akin to the "attractions" of early cinema—views of exotic places, reports of news events, short animations, sensations like frightening experiences, and so on—but there is nothing that resembles the kind of sustained storytelling that contemporary moviegoers enjoy, either fictional or nonfictional. With.in's own studio has turned its attention from VR storytelling to producing VR workout videos.

To VR skeptics, including myself, it comes as no surprise that the new medium has not taken off as a vehicle for storytelling. One obvious reason is the headset. It is inconvenient and uncomfortable. The lag between head movement and the corresponding scan of the virtual environment in the viewer, although barely perceptible, gives some people nausea or headaches. And wearing the headset gives people the feeling of being cut off from the real world, including other people, which is particularly ironic with respect to documentary, which is supposed to do just the opposite.

On the other hand, VR is taking off as a platform for gaming. VR gamers are evidently not troubled by the discomfort or sense of isolation caused by the headset or, if they are, they deem it a price worth paying for the immersive experience. This shows that the headset isn't the main reason for the lack of success of long-form VR stories. Something else is to blame. I maintain that that something is, simply, the fact that the VR medium is not conducive to telling stories. Certain intrinsic qualities of the medium actually get in the way.

Think about what stories are and how they work. Stories don't show you characters doing random things; they show you characters doing things *for particular reasons*. It isn't the actions that are central; it is those reasons. Stories work by inviting and allowing you to "get into characters' heads," so to speak: to imagine their intentions, desires, emotions, beliefs, knowledge, and so on. Psychologists commonly refer to this sort of perspective taking as "theory of mind." That is how we navigate the social world. Most of the time, we do it automatically, without thinking about it. We can also do it deliberately, when we want to explain or judge

somebody's behavior, for instance. And we can do it just for fun, as when we read novels and watch fiction films. Trying to make sense of the *why* of people's actions and behavior is engrossing and satisfying. That is the beating heart of all stories and storytelling.

Novels and spoken stories help us make sense of why characters do what they do by describing their thoughts and feelings expressly, with words. Pictorial media, like paintings and movies, depend largely upon inference, instead, since it is impossible to directly depict what is going on in somebody's head. But cinema has tools that can guide such inferences extremely effectively. With its frame, it can foreground important clues, like close-ups of characters' facial expressions. The frame also allows for control of viewers' vantage point. Plus, along with editing, the frame permits cinematic storytellers to minimize or simply sideline anything that is not directly relevant to making sense of characters' actions. That is why cinema is such an effective storytelling medium.

VR has no frame. Its effect is to put you in the middle of a space or event. The only way VR can create the equivalent of cinema's facial close-up is by having a character or subject approach very close to the camera. In VR, this feels looming and weird. The only way VR can keep viewers from looking around to see what may be off to the side or behind them is by blurring it or blacking it out. This defeats the whole purpose of VR, which is to give viewers the freedom to look around. VR sometimes uses movement, light, and sound cues to call viewers' attention to particular elements of a scene, but those tend to be weak clues as to what is going on in characters' heads. Precisely because VR lacks a frame, it is an ineffective storytelling instrument when compared to conventional cinema. But there is an even bigger problem.

The very fact that you can look around in 360 VR is an *invitation* to look around. The very fact that you can interact with elements of the environment in VR games is an *invitation* to interact. In the first case, your attention is directed toward the environment; in the second, it is directed toward the gameplay. In both cases, the focus turns to your *own* experience, as opposed to the experience of other people in the space. This can be interesting, worthwhile, and fun, but instead of helping you understand what characters or subjects in the scene are thinking and feeling, which is what is most important from a storytelling perspective, it tends to interfere with or distract from it.

I'm discussing the experiential dimension of VR here. Because this is impossible to fully convey in writing, I refer you to a video essay I authored, with concrete examples and illustrations, entitled "Why VR Does Not Promote Empathy." It can be watched online at *[in]Transition: The Journal of Videographic Film & Moving Image Studies*, 7.2, 2020.

Incidentally, whether these experiential differences have any measurable impact with respect to viewer prejudice and pro-social behavior is debatable. A very interesting empirical study by Harry Farmer designed to test just this, using *Clouds over Sidra*, found no significant difference between VR, a flat-screen version of the movie, and a written transcription. This study is reported in *Immerse* under the title "A Broken Empathy Machine?" (2019). On the other hand, a study led by Fernanda Herrera at Jeremy Bailenson's VR lab at Stanford, comparing a VR and first-person written account of the experience of becoming homeless, found that the VR experience had a significantly stronger and longer-lasting effect in prompting participants to sign a petition in support of affordable housing (2018). The debate about whether documentaries of any kind "really make a difference" is an old and ongoing one. They clearly make a difference *to viewers*. I would argue that that is sufficient. In any case, that is what I am discussing here.

Early theorists of VR, including documentary scholar William Uricchio, anticipated my reservations about the medium and respond the them. Look, he wrote in 2016, "VR is not film":

> Just as film's pioneers spent their first decade emulating theater, many of today's VR makers are doing their best to emulate the logic of film. And just as early filmmakers finally shattered the proscenium arch and evolved new vocabularies, we can expect VR's makers to find robust, exploration-based alternatives to the still-dominant film paradigm of carefully composed frames, editing strategies and fades to black.

VR's only problem, in other words, is that it has not yet discovered or invented the right means for effective storytelling.

In a response to my video essay, new-media scholar Miriam Ross raises another objection. In conceiving storytelling as I do, in cinematic terms, she says, I am conceiving stories and storytelling too narrowly. There is not just one kind of story, she suggests; there are many. VR pioneers are not just figuring out how to tell conventional stories with new and different means, they are exploring new forms of story. In so doing, she implies, they may be helping to expand our very idea of what stories and storytelling mean.

A student of mine, an avid video-gamer, made the same point to me more concretely and forcefully. She told me, imagine someone arguing, in the early days of video games, that video games will never succeed as a vehicle for storytelling because they do not tell stories in the same way that movies do. That argument has been proved dead wrong. The video game industry now makes more money than the movie and sports industries combined. And video games do in fact tell stories. Video-game stories are different from the kind of stories that movies tell but they are stories just the same. In video-game stories, instead of imagining other peoples' stories, as in movies, you *experience* stories yourself. You are, in effect, the protagonist. At the same time, you take on the role of an imaginary character, who may be nothing like the actual you, and you encounter situations and events that are nothing like those you experience in real life. In this way, my student argued, video games may be even better at mobilizing and developing perspective-taking skills than cinema is. If today's VR documentaries fall short as stories, the reason is not that they haven't discovered the right techniques; it is that they haven't discovered the right kind of story.

In the early days of academic video-game theory, a debate raged between "narratologists," who argued that video games did indeed tell stories, albeit of a new form, and "ludologists," who argued that what is most interesting and distinctive about video games had nothing to do with stories and storytelling. An essay by Harmut Koenitz in the *Encyclopedia of Computer Graphics and Games* (2018) gives a very nice overview of this debate. It gradually settled down as it became clear that the difference between the two camps had less to do with video games *per se* than it did with their definitions of narrative. If children playing with toys, football games on TV, and scientific papers can all be usefully thought of as forms of stories and storytelling, then surely video games can, too. (The narratologists' view.) If, on the other hand, one thinks about these other forms of discourse in terms of stories and storytelling, that tends to flattens out the very significant differences between them. (The ludologists' view.) Both views have merit. It is worthwhile to think about both the similarities and the differences between video games and conventionally structured narratives. In the same way, it is worthwhile to think about both the similarities and the differences between VR narratives and cinematic narratives. But to do this fruitfully, we need to start by agreeing on a definition of narrative.

2   What Is a Story, Anyway?

The broadest, most inclusive definition of story probably goes something like this: Story is the means by which we understand and explain human behavior. It is also the means by which we navigate and imaginatively engage in the social world. As such, it is an intrinsic part of all social life. For example, when you watch even the most abstract experimental film, you can hardly help but "narrativize" it. When you wonder "What does this *mean*?" or when you ask yourself, "What does the filmmaker *intend* by this?" you are thinking about the movie in terms of story.

Telling stories intentionally invites this kind of response. It is the use of some form of discourse, such as speech or cinema, *for the purpose of* prompting people to imagine the mental and social lives of other people. An abstract experimental film might be designed solely to prompt an aesthetic experience. In that case, it is not storytelling; it is "art" or something else. But if part of its intended purpose is to prompt viewers to reflect on the lives and experiences of themselves or other people, as appears to be the case in the semi-documentary works of Stan Brakhage, for example, then it can be regarded as a form of storytelling, even if the stories it elicits are relatively undirected and open to viewers' subjective interpretations.

By this expansive definition, most VR movies are definitely a form of storytelling.

In view of this fact, I need to tweak the way I describe the shortcomings of VR. Its problem is not that it is an intrinsically weak storytelling instrument; its problem is that, rather than focusing our attention on the mental

and social lives of *other* people, in the way that movies typically do, it almost inevitably draws our attention toward our *own* experience, much like video games.

My video-gaming student critic might ask, So what's wrong with that? Part of what's great about video games is that they allow you to have your own adventures, to make your own choices, to experience things for yourself. If VR does that, too, isn't that a totally good thing?

Sometimes, but sometimes not. It depends upon the purpose of the VR movie. If the purpose is make-believe play—to imagine piloting a spaceship or fighting off zombie attackers—then by all means have your own experience and enjoy it. The same is true if the purpose of a VR movie is to be evocative or prompt an aesthetic experience, like the films of Stan Brakhage. In that case, the effect of being able to look around could be marvelous. So, too, if the purpose is to simulate the experience of swimming with sharks off the Great Barrier Reef (although I would call this simulation, not storytelling).

But what is the purpose of virtually visiting a Syrian refugee camp? Is it to explore? To have an adventure? To have a stimulating aesthetic experience? Or is it to genuinely understand the mental and social lives of the people who are forced to live there? If that is the goal, then paying attention to your *own* experience will definitely get in the way.

I believe that one of the most fundamental and worthwhile purposes of storytelling is to help us imagine other persons' experience. Advocates of VR "empathy" seem to believe this, too. Their mistake is to suppose that just being in the same place as another person, actually or virtually, automatically confers a deeper understanding of that person's mental and social life. It *may* confer deeper understanding, but only if our focus remains firmly on the mental and social life of that other person. VR's perpetual invitation to look around is a constant distraction from this. That's my beef with VR as a storytelling instrument.

Nevertheless, there is real emotional power in being able to look around in VR, to be free of the dictates of cinema's frame, to "see for yourself" as it were. So, it may be, as my video-gaming student critic argues, that there are ways of prompting or inviting us to imagine the mental and social lives of other people that are different from cinema and better suited to VR. What might those look like?

Mainstream movies are designed to so engross us in the lives and experiences of characters that we forget ourselves. Media scholars sometimes refer to this phenomenon as "transportation," following literary theorist Richard Gerrig (1993) and media psychologists Melanie Green and Timothy Brock (2002). Movies accomplish it by foregrounding characters' problems, desires, and emotions. Extraneous or background elements, like visual effects and music, largely serve to shape or reinforce our understanding of what characters are thinking and feeling. All of our attention and imagination is focused on the events we see unfolding on the screen and, in particular, on the impact those events have on the characters. Even though we may wind up "identifying" with those characters, we are just observers. Even though we may project our own feelings on the characters and compare our own experiences to theirs, the story we are experiencing is always *their* story, not ours. For that reason, we can think of this as third-person storytelling.

This is an admittedly over-simple description of the way movies work—especially documentaries, as I will explain later—but it gets to the heart of what makes "escapist" Hollywood movies so popular.

Video games can induce a state of absorption similar to narrative transportation. Game theorists refer to this as "flow"—a term borrowed from psychologist Mihaly Csikszentmihalyi (2008). But what puts you in this state is not a focus on the problems, desires, and emotions of characters, as with movies; it is on dealing with the challenges of the game. Even though, in most video games, the experience involves taking on the role and even the body of a virtual character, and even though your character's task is typically depicted in social terms, like making money or vanquishing enemies, what puts you in that state of flow is not imagining what characters in the game may be thinking and feeling; it is mastering the skills you need to advance to the next level. Even in the case of a video game like *The Last of Us*, which is famous for making players cry on account of its bittersweet narrative ending, the feeling of sadness is mostly relevant as part of the player's experience. The player's experience is the primary focus. For that reason, we can think of this as first-person storytelling.

This description of video-game play is, again, over-simple. In fact, the emotionally moving scenes in *The Last of Us* and similar games are cinematic "cutscenes" in which the player is a mere observer and has no control.

Third-person scenes, in other words. First-person applies mainly to the gameplay sequences. In *The Last of Us*, these gameplay sequences revolve around the relatively simple and repetitive actions of avoiding or dispatching zombie-like creatures and human enemies. These actions are the heart of the gameplay experience and are what put gamers into that state of flow. They are obviously a very long way from imagining the thoughts and feelings of other people.

Although I'm using the terms first- and third-person in a somewhat idiosyncratic way here—to refer to two different kinds of experience, instead of two different kinds of narration, as film and literary scholars typically do—there is nothing particular new or controversial about the difference I am describing between the experience of movies and the experience of video games. But there is a fascinating social dimension to this difference that is not often discussed. When you watch a movie, because what matters is somebody else's experience, you are generally supposed to silence your cell phone, shut your mouth, and pay attention. If you feel a social or emotional connection to somebody during a movie, it is most likely to a person on the screen. Video games are very different in this regard, as is especially apparent from online games like *Fortnite*. Because the focus of video-game play is your own experience, you are more than welcome to share it. Through such sharing, "my" experience often becomes "our" experience—still first-person, but first-person plural. This can create the feeling of a social and emotional connection to other players, especially those on your team. The downside is that other characters, including folks on the other team, can be effectively reduced to incidental props or opponents to be beaten. So, the difference between what I'm referring to as first- and third-person storytelling clearly has important ramifications if your goal is to use stories to foster caring or compassion for others.

I argued earlier that, like video games, VR stories constantly prompt you to attend to your own experience. But that's not the whole picture. Video games in VR are, for all intents and purposes, just regular video games, with the addition of stereoscopic vision and the use of the headset rather than a thumb controller to look around. In other words, they are first-person stories in exactly the same way traditional video games are, if a bit more lifelike or vivid. But with VR documentaries, like *Clouds over Sidra*, something more complicated is going on. The stories they tell are not just experienced in third person, in the way of entertainment films, and not just in first person, as with video gameplay. They have a second-person quality.

In second-person discourse, you have a sense of being in a direct relationship with someone. When somebody writes you a letter, asks you to pass the salt, or "likes" one of your Facebook posts, that is second-person discourse. In a 1975 essay, film semiotician Christian Metz distinguished "two kinds of voyeurism": story and discourse. In story, he said, you are not aware of the presence of an author or interlocutor; in discourse, you are. Metz proposed a psychoanalytic explanation for the invisibility of the author in mainstream movies. A much simpler cognitive-psychological explanation is that we just tend not to notice what we are not paying attention to. (You can see some experiments that dramatically and sometimes humorously demonstrate this psychological phenomenon by searching for "inattentional blindness" on YouTube.) But Metz's distinction remains useful. The difference between what he calls story and discourse is the essential difference between what I'm referring to as third-person and second-person storytelling.

In VR documentaries, because of VR's power of immersion, you have a sense of being in the same space as subjects, even though you cannot interact with them. This effect is particularly pronounced when they look or speak directly to you (to the camera, actually), something that rarely happens in fiction films, precisely because it disrupts the "transparence" of third-person storytelling. Our attention, in these moments, is subtly drawn to our relationship with the subject. It is important to note that this is not the same thing as focusing on or prompting us to imagine what that person is thinking or feeling, the way third-person storytelling does. Still, the visceral sense of the subject's "presence" might facilitate our thinking about and even caring about what the person who is virtually next to us is thinking and feeling—in a different way from a cinematic close-up, obviously, but perhaps to similar effect.

Some of the most compelling experiments in VR fictional storytelling make deliberate use of this second-person effect. An excellent example is *Wolves in the Walls*, which won an Emmy in 2019 for outstanding innovation. The story, based on a children's book by Neil Gaiman, is about a little girl who hears scratching in the walls and, because nobody believes her, sets out to prove that something is really there. What is most fascinating and engaging about the way the story is told is not "the feeling of being there" in VR; it is way the

little girl constantly interacts with you. She looks directly at you when she talks and her eyes follow you when you move about. If you try to duck or dodge out of her line of sight, she asks, "What are you doing?" She hands you things, like a virtual camera, and invites you to do things with her, like draw on the wall. When you observe something in the story in third-person mode, like the girl's mom canning in the kitchen, the girl is by your side, looking with you and talking to you. This is not a video game, where the focus is on your own choices. The focus, most of the time, is on the girl and her story. But you are not just an observer, either. Instead, you are part of the story, as the girl's friend and confidante. *Wolves in the Walls* is an outstanding example of effective second-person storytelling.

From this example, you might suppose that effective second-person storytelling with movies requires interaction, in much the same way that effective third-person storytelling with movies requires a frame. But, as it happens, there is a popular form of conventional cinema that routinely and systematically uses forms of second-person storytelling. Documentary!

Conventional cinematic documentaries are full of explicit second-person discourse, like interviews and voice-over narration. There is also an implicit sense in which documentaries *are* second-person discourses, in their entirety. Like allegorical fables and newspaper editorials and "thesis" novels, we know they are supposed to have a message or point. They are trying to "tell us something."

This suggests that there may be a natural affinity between VR as a medium and documentary storytelling, having to do with second-person address. It also suggests that, in spite of the shortcomings of VR as a means of understanding what is going on in characters' or subjects' heads, VR may have compensatory advantages for documentary storytelling, having to do with the sense it can give of being in a relationship with other people by virtually putting you in the same space as them.

3   Second-Person Storytelling in VR Documentary

In his popular textbook *Introduction to Documentary*, Bill Nichols describes the many different ways in which documentaries "speak" to us. For example, a documentary may imply, "I speak about *them* (other people) to *you*." Or it may suggest, "I speak about *us* to *you*." But "you" is always part of the equation. As Nichols writes,

> Without this sense of active address, we may be present at but not attend *to* the film. [Documentary] filmmakers must find a way to activate our sense of ourselves both as the one to whom the filmmaker speaks (about someone or something else) and as members of a group or collectivity, an audience for whom this topic bears importance (44).

Nichols is describing documentaries' implicit second-person address.

Usually, the implicit purpose *of* a documentary and the explicit second-person address *in* the documentary are closely aligned. For example, with *March of the Penguins* (2005), Morgan Freeman's sonorous narration, the film's awesome aerial shots, and its personification of penguins all fit perfectly with the movie's primary aim as a documentary, which is to foster appreciation of nature. Even though the techniques work to draw us into the story, they are also constant subtle reminders of the overarching purpose of the story. This is not just a regular entertainment film, they say; it is *important*, since it is about the life-and-death struggle of actual penguins. In the documentaries of Michael Moore, the alignment is even more direct. Since Moore is the main character in his films as well as the filmmaker, he more or less tells us, outright, what we are supposed to think (whether or not we happen to agree with him).

In VR documentaries like *Clouds over Sidra*, in contrast, there is a subtle misalignment between the way the movie "speaks" to us and its purpose. The maker of the movie, Chris Milk, has told us what the purpose of the movie is: it is supposed to create "empathy" with its subject, a 12-year-old Syrian girl in a refugee camp in Jordan. The movie does succeed in giving us a vivid and visceral sense of her physical situation: rows of tents in a filthy grassless environment, the living room of the girl's family's tiny apartment, her schoolyard, surrounded by a barbed-wire fence, and so on. But, as I've said, it tends to distract us from the girl herself. The effect is to prompt us to think (or feel), "How would *I* like to live in a place like that?" or, sometimes, "I wonder what I will see if I turn around."

Is there any harm in this? Or, to put the question positively, might there be some benefit of this effect, with respect to capturing viewers' attention, prompting sympathetic emotional responses, and even fostering genuine understanding of and compassion for other people? That is the important question, not whether we are focusing on the girl or her surroundings.

The second-person address of documentaries can elicit gratifying social emotions. That is a fundamental part of the appeal of documentaries and of nonfiction generally, and part of what we get from them in exchange for the fact that, as stories go, they are usually comparatively dull.

One of the social-emotional gratifications of documentaries is the appeal of acquiring knowledge or edification. This is basically an appeal to our egos. In watching a documentary, we may imagine ourselves to be especially smart and sophisticated, morally superior, privy to special knowledge or inside information, or part of a community that shares particularly worthwhile values, beliefs, or concerns. A second social-emotional appeal is the sense of an implicit alliance with the filmmaker, involving reciprocity: the filmmaker owes you honesty and you owe the filmmaker trust. With respect to these two gratifications, VR documentaries like *Clouds over Sidra* are no different from conventional documentaries. But consider a third social-emotional appeal.

Documentaries appeal to us as social actors: inhabitants of a shared social world, with responsibilities and concerns that go along with that. "Appeal" here implies both an attraction (something engaging and gratifying) and an entreaty (a request for attention and possibly action). In documentaries these two kinds of appeal are related. The second-person addresses of documentaries give us a sense of being recognized and acknowledged. I believe this is what Bill Nichols means when he says that documentaries "speak" to us and "activate our sense of ourselves." This kind of appeal is not specific to documentaries. It is implicit in any second-person discourse, whether that is "please pass the salt," Morgan Freeman telling you about penguins, or the little girl in *Wolves in the Walls* following you with her eyes.

In the case of Morgan Freeman telling you about penguins, the appeal is indirect and not particularly strong. It requires some investment on your part. To help create this investment, *March of the Penguins* leans on dramatic stories and cuddly penguins. The "story" in the VR documentary *Clouds over Sidra* is not nearly so intrinsically engaging—it consists largely of the refugee girl telling you about how she spends her days. On the other hand, there is something undeniably compelling about the girl sitting across from you in VR and speaking to you directly. "The feeling of being there," in the same space as the girl, gives her address to you a kind of immediacy and force. The perpetual temptation in VR to gawk around, to see what is behind you or off to the side, is modified by a subtle emotional pressure to pay attention to the girl, prompted by the sense of her "presence" (a word very commonly used in discussions of VR). Notably, this emotional pressure is absent when the girl is not on screen.

Consider the social ramifications of this kind of address, compared to those of movies and video games. The "transparent" third-person address of mainstream fiction films creates an imaginary connection to characters in the movies. You "put yourself in their shoes," so to speak. That is the power and the appeal of that kind of storytelling. The first-person address of video-game storytelling invites you to focus on your own experience. That is its power and appeal. But you can also share your experience with others, which creates the feeling of a social bond or connection. Second-person discourse has a similar effect, of creating the feeling of being in a relationship with someone. In fact, the chatter that takes place during online multiplayer video games *is* second-person discourse.

What's special about the second-person addresses in VR stories is the sense of being *with* somebody else, as a companion or associate or friend. This creates a subtle feeling of obligation toward the person speaking to you. You *could* look away, because this is VR, but perhaps you *shouldn't*. This sense is by no means the same thing as empathy. Still, it may be a stepping stone to *caring*. So, if VR has some special strength for documentary storytelling, I think it is its capacity to virtually put you in the presence of other people, who can speak to you directly. VR is not nearly as good as conventional cinema for creating understanding of others, because if its lack of a frame, but perhaps in some cases it can compensate for or even overcome that weakness by facilitating a sense of social connection with and obligation toward its on-screen subjects.

Other VR researchers and practitioners have come to a similar conclusion, from other angles. A notable example is a 2018 article on the ethics of VR documentary, by documentary scholar Kate Nash, in which she discusses "proper" and "improper" distance. She writes, "VR runs the risk of producing improper distance and an ironic mode of moral engagement when it invites forms of self-focus and self-projection rather than a more distanced position that allows for recognition of distance between the self and other." She's describing here what I call first-person storytelling. It's fine for video games, as I've said, but where understanding others is the goal, Nash argues, it is morally problematic.

Nash's examples of "proper" distance in VR all involve what I refer to as second-person storytelling. In her words, these works "do not simulate the other's experience or aim to put the user in their shoes directly. Rather, they simulate an encounter, a meeting with the other" (126). One of her examples is a scene from *My Mother's Wing* (2016), a VR documentary about a Palestinian woman trying to cope with the loss of her two children who were killed in a shelling attack. We are positioned as though in conversation with the mother, sitting on the floor, as she addresses her two lost children. The VR makes this feel like a face-to-face encounter. That in turn creates a sense of immediacy and moral responsibility. But even in moving moments like this one, the VR medium is also pulling us in another direction, Nash observes:

> In these moments of simulated face-to-face encounter, the user is also being invited to explore a new space.... The effect is to encourage a profound turning away from the speaking subject. The user's attention is divided but more than that, the physical turning away that visual exploration demands is profoundly at odds with the moral demand of the face-to-face encounter. There is a risk of improper distance insofar as the user prioritizes their own experience of transportation and exploration (spatial presence) over engagement with the testimony of the other. Similarly, insofar as the user focuses on their experience of transportation as indicative of 'what it is like' to be in the space of suffering it is possible to speak of improper distance (128).

Nash's argument is conceptual. She does not discuss the practical implications of what she calls improper distance. But there is empirical research on the same phenomenon in real-world simulations, like wheelchair-for-a-day experiences, that sheds light on its social and psychological consequences. Michelle Nario-Redmond and her co-researchers review this research in a 2017 article, "Crip for a Day." Although disability simulations are intended to promote empathy in much the same way that VR documentaries are supposed to, the research shows that their actual effects are more complicated. On the positive side, they engage people, stimulate discussion, and generate warm feelings and altruistic impulses toward disabled people. But they have dangerous unintended consequences, as well. They tend to result in pity and guilt, self-orientation and smugness, frustration and fatigue, and a corresponding tendency to avoid people with disabilities. On the whole, some researchers conclude, disability simulations "serve to constitute and reproduce, rather than disrupt, disability oppression." In a similar way, when you drop in on a refugee camp for a few minutes via VR, there is the danger that you are involved in a kind of virtual voyeurism that results more in feelings of self-righteousness and pity than in genuine understanding and compassion.

4 Practical Take-Aways

There is social and emotional value in visiting a refugee camp. There is value in being present at a culturally significant event, like a political rally. There is value in having an unusual experience, like swimming with sharks. There is no doubt also genuine social and emotional value in experiencing such things as these virtually, through VR. So, VR is without question a valuable medium for nonfictional representations and discourses. *But* if your priority is to foster genuine understanding of other people's lives, hopes, fears, pains, happiness, and so on, then you are better off using conventional cinema, since that is better than VR at keeping your audience's attention focused on other people and more effective than VR at conveying other people's thoughts and feelings. In other words, for third-person storytelling, conventional cinema simply works better than VR does.

Here, I need to insert a tangential clarification. Throughout this essay, I have been discussing third, first, and second-person experiences as though they are separate and discrete. In fact, they are regularly intermingled, in fiction films, video games, documentaries, and VR experiences. So, it might be best to think of them as the

center of narrative gravity of moments or scenes. "Center of narrative gravity" is actually a term philosopher Daniel Dennett (1992) coined to describe his conception of the self in consciousness. I use the same term here because I find it quite helpful to think about the three modes of narrative experience as three different modes of self-consciousness. Our creative imagination can focus on three different things when we think about social situations: our own subjective experience (first person), our understanding of the experience of others (third person), or our relationship with someone else (second person). We can only focus on one of these at a time, because of the limits of attention. Nevertheless, when watching VR documentaries, we experience a kind of tension between these perspectives and switch back and forth between them, in the way I have described.

Anyway, what VR is really good at, much better than conventional cinema, is first-person storytelling: helping us imagine viscerally, through firsthand experience, what it feels like to live in a refugee camp, to attend a political rally, or to swim with sharks. VR simulations of such experience are interesting and worthwhile in their own right and can be extremely useful for reporting on some situation or event. Where empathy is concerned, however, this usefulness comes with a caveat. When we're focused on our own experience, it is hard to genuinely imagine somebody else's. We may *think* we are imagining somebody else's experience but, as with wheelchair-for-a-day experiences, our self-absorption tends to eclipse genuine understanding. Note, however: this weakness by no means precludes genuine *caring* for others.

All of the discussion about whether or not VR fosters empathy obscures the very important difference between empathy and caring. Think about the difference this way: you can truly and deeply care about the death of coral reefs without empathizing with coral reefs. In the same way, you can truly and deeply care about the plight of Syrian refugees without empathizing with any particular refugee. Third-person stories (e.g., stories about the plight of particular Syrian refugees) definitely facilitate understanding; understanding naturally contributes to caring; and VR is unfortunately not very good at third-person storytelling. Just the same, there are other means of fostering caring that VR is able to use quite well. Three, in particular: creating interest, conveying information about places, and cultivating a sense of social connection.

"The feeling of being there" in VR is intrinsically interesting, especially if the place you are virtually visiting is foreign or spectacular. This interest can be leveraged into interest about a topic, whether that is coral reefs or the plight of refugees. It does not require empathy; all it takes is attention. Getting somebody to put on a VR headset is surely a way of capturing their attention. Unfortunately, this kind of interest does not last very long. As soon as the novelty wears off, people want more, like a story or a game. Stories and games are what sustain viewers' interest. For telling stories, conventional cinema works much better than VR.

VR does have the unique ability to convey information about a place or environment. Words and two-dimensional images cannot convey what it looks like from the top of Mount Everest. VR can. In the same way, VR has a unique ability to convey what happens to a coral reef when it dies or what life in a refugee camp is like. Knowing what life in a refugee camp is like does not create empathy, any more than knowing what happens when a coral reef dies, but both can surely contribute to understanding, and that can lead to caring. So, if information about a place or situation is central to a documentary story, VR is an excellent means of conveying that. The traveling shots through erstwhile Nazi concentration camps in the documentaries *Night and Fog* (1956) and *Shoah* (1985) come to mind as examples of scenes to which VR could have powerfully contributed. But, again, if understanding the experience of other people is what is most important, you are better off with traditional cinema.

Finally, VR has a unique ability to create a sense of the "presence" of another person, which invites attention and respect for that person. As with interest generated by the novelty of "the feeling of being there," this effect is transitory, not enough to hold viewers' attention for very long. And, as with VR's information about place or environment, it does not result in empathy. But it does tend to generate a feeling of moral obligation, as Kate Nash observes, and that is an essential ingredient of caring.

There is a famous interview in *Shoah* in which Abraham Bomba, a man who had been forced to work as a barber in a concentration camp, breaks down while describing one of his barber friends coming across his own wife and daughter while cutting hair in the anteroom of a gas chamber. The scene is shot very simply: Bomba sits at a table, framed in a medium shot. As a thought experiment, I tried to imagine what this scene would feel like had it been shot in VR. On one hand, VR would likely have weakened the scene, even cheapened

it, by inviting me to look around at the space. On the other hand, if I'd had the feeling of sitting at the table across from Bomba, I would have found it very difficult to look away. This is the moral ambivalence of VR that Kate Nash describes.

There is also a third possibility, which is that with VR I'd have been able to look back and forth between Bomba and his interviewer, filmmaker Claude Lanzmann. That would have created a different kind of social and situational awareness. I suppose this would have diluted my attention to Bomba's story and made me less likely to cry. On the other hand, I suppose it would have allowed me to experience a deeper and more genuine empathy with Bomba who, in this moment, is being asked to revisit horrors of his past for the viewers' sake. I would have had a fuller, truer picture of the second-person discourses in which I was involved. There are VR documentary filmmakers who have called for more of this kind of transparency in VR. Instead of literally hiding, as the camera operators of *Clouds over Sidra* and other VR documentaries do, they encourage VR filmmakers to include themselves and their tripods in their scenes. This would have the advantage of taking the second social-emotional appeal of documentaries that I mentioned earlier—the implicit contract between filmmaker and viewer, involving honesty and trust—and putting it right in the movie, without expressly calling attention to it or foregrounding it in the way a conventional documentary would need to do, by cutting to shots of the filmmaker.

There is one other more interesting creative possibility, which is to abandon the goal of quasi-cinematic storytelling altogether and instead use the more game-like qualities of interactive VR to try to create a sense of social connection with documentary subjects. The *Book of Distance* (2020), by artist and filmmaker Randall Okita, is an interactive animated VR documentary that succeeds quite well in doing this. The subject of the piece is the artist, himself, as he tries to imagine the experiences of his Japanese grandfather, who immigrated to Canada just before World War II. These imaginings are presented in animated vignettes, which often include family photographs and other historical artifacts that can be interacted with. The effect is to engage viewers in the artist's personal quest to make sense of his family history.

The question is, might this have all been done more effectively by conventional cinematic means. The VR medium provides an interesting, compelling, first-person experience. But does it ultimately bring us any deeper understanding of or genuine caring for Okita and his grandfather? Or does it mainly make us feel satisfied for self-centered reasons, like a video game?

This question calls to mind a 2017 essay in *The Atlantic* by video-game theorist Ian Bogosian, entitled "Video Games are Better without Stories," in which he discusses the video game *What Remains of Edith Finch?* The mechanics of that game are similar to the mechanics of *The Book of Distance*. You (the player) are Edith Finch, the last survivor from a cursed family, exploring a vacant house you inherited from your deceased mother. As you explore, two things happen: first (as in *The Book of Distance*), you gradually uncover stories about your doomed ancestors, through diaries, recordings, and mementos; second (unlike *The Book of Distance*), you encounter weird new spaces in the house and experience weird altered states, like taking on the form of animals. Bogosian found the game satisfying. Still, he muses, "Why does this story need to be told as a video game? The whole way through, I found myself wondering why I couldn't experience Edith Finch as a traditional time-based narrative." Bogosian's answer, essentially, is that because then he would not have had all of those weird and wonderful aesthetic experiences. In other words, the game's "story" succeeds because it abandons the whole idea of third-person storytelling. It is not really about a person named Edith Finch. It really is about *you*, the player. And that's good, Bogosian seems to conclude, because *What Remains of Edith Finch?* is, after all, a game.

But what if Edith Finch's ancestors were real people, with real pasts and real stories. In that case, treating her past as something to "explore" and "experience" game-wise might be giving short-shrift to those people's actual lives and experiences. What saves me from such misgivings, with respect to *The Book of Distance*, is that the real subject of the documentary is the author himself. He speaks to us directly, in his own voice, throughout the experience. So, in our first-person explorations and experiences, we have a sense of partnering with him.

To return to my earlier question, could this have been done more effectively with conventional cinematic means? The answer is no. That might well have resulted in a more directed, more focused, possibly even a more moving story about Okina's grandfather, but it would have sacrificed much of the sense of partnering

with the artist in an exploration of his family history. So, *The Book of Distance* is one example of documentary storytelling that is more effective in VR than in conventional cinema. But notice that *The Book of Distance* does not ask us to "identify" with Okina's grandfather or to "empathize" with him or be "transported" into his life and experiences. Instead, we are invited to remember him and to reflect on events in his life by experiencing his grandson's artistic re-imaginings of them. That is second-person storytelling.

If the goal of a documentary is third-person storytelling, however—that is, to provide a kind of window into the life, thoughts, and feelings of somebody other than the filmmaker and the viewers (such as, for example, a Syrian girl living in a refugee camp), then you are better off with conventional cinema. To put it more bluntly, *Clouds over Sidra* is an interesting and briefly engaging VR experience, but it is not a particularly good documentary. Nevertheless, because it uses VR, it may prompt some people to put on a headset and pay attention, for a few minutes, to what it looks and feels like to be in the middle of a refugee camp. Even though it is a mistake to suppose that this creates compassion, by itself, it can help to lay a groundwork for compassion by drawing interest to and focusing attention on the situations of other people who are in need not just of compassion, but also of political advocacy and material support, like the countless refugees in the U.S., Europe, and around the world. And that is, without question, a good thing.

References

Bloom, Paul (2017). "It's Ridiculous to Use Virtual Reality to Empathize with Refugees." *The Atlantic*, Feb. 3. https://www.theatlantic.com/technology/archive/2017/02/virtual-reality-wont-make-you-more-empathetic/515511/.

Bogosian, Ian (2017). "Video Games are Better without Stories." *The Atlantic*, April 25. https://www.theatlantic.com/technology/archive/2017/04/video-games-stories/524148/.

Csikszentmihalyi, Mihaly (2008). *Flow: The Psychology of Optimal Experience*. New York: Harper Perennial Modern Classics.

Dennett, Daniel (1992). *Consciousness Explained*. New York: Back Bay Books.

Farmer, Harry (2019). "A Broken Empathy Machine?" *Immerse: Creative Discussion of Emerging Nonfiction Storytelling*, Sept. 30. https://immerse.news/a-broken-empathy-machine-can-virtual-reality-increase-pro-social-behaviour-and-reduce-prejudice-cbcefb30525b.

Gerrig, Richard J. (1993). Experiencing Narrative Worlds: On the Psychological Activities of Reading. New Haven, CT: Yale University Press.

Green, M. C. and T. C. Brock. "In the Mind's Eye: Transportation-Imagery Model of Narrative Persuasion." In *Narrative Impact: Social and Cognitive Foundations*, edited by Green, M.C., J. J. Strange and T. C. Brock, 315-41. Mahwah, NJ: Lawrence Erlbaum.

Herrera, F., J. N. Bailenson, E. Weisz, E. Ogle, and J. Zaki (2018). "Building Long-Term Empathy: A Large-Scale Comparison of Traditional and Virtual Reality Perspective-Taking. *PLoS ONE* 13(10): e0204494. https://doi.org/10.1371/journal.pone.020449.

Koenitz, Helmut (2018). "Narrative in Video Games." In *Encyclopedia of Computer Graphics and Games*, Living Edition, edited by Newton Lee. New York: Springer Link. https://doi.org/10.1007/978-3-319-08234-9_154-1.

Metz, Christian (1981). "Story/Discourse (A Note on Two Kinds of Voyeurism)." In *The Imaginary Signifier: Psychoanalysis and the Cinema*, 89-98. Bloomington: Indiana University Press.

Milk, Chris (2015). "How Virtual Reality Can Create the Ultimate Empathy Machine." TED: Ideas Worth Spreading. https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine.

Nario-Redmond, Michelle, Dobromir Gospodinov, and Angela Cobb (2017). "Crip for a Day: The Unintended Negative Consequences of Disability Simulations." *Rehabilitation Psychology* 62.3: 324-333. https://pubmed.ncbi.nlm.nih.gov/28287757/.

Nash, Kate (2018). "Virtual Reality Witness: Exploring the Ethics of Mediated Presence." *Studies in Documentary Film* 12.2: 119-1301. https://doi.org/10.1080/17503280.2017.1340796.

Nichols, Bill (2018). *Introduction to Documentary*, 3rd ed. Bloomington: Indiana University Press.

Uricchio, William (2016). "VR is Not Film, So What Is It?" *Immerse: Creative Discussion of Emerging Nonfiction Storytelling*, Nov. 8. https://immerse.news/vr-is-not-film-so-what-is-it-36d58e59c030.

Dirk Eitzen – Franklin & Marshall College (United States)

✉ deitzen@fandm.edu

Dirk Eitzen is Professor of Film & Media Studies at Franklin & Marshall College in Lancaster, Pennsylvania. He has produced a dozen documentaries, including The Amish & Us (1998) and Save Our Towns, Save Our Land (2000), which both aired nationally on U.S. public television. His scholarly work includes numerous theoretical essays on documentary including, most recently, "The Duties of Documentary in a Post-Truth Society" (in Cognitive Theory and Documentary Film , eds., Catalin Brylla and Mette Kramer, 2018).