Exploring the revitalisation of endangered intangible heritage and languages through multimedia storytelling and immersive technologies: a case study of virtual reality and 2D film with the Kusunda community in Nepal

Fabrizio Galeazzi [1]*, Gayatri Parameswaran[2], Felix Gaedtke[2] and Viren Swami [3,4]

[1]StoryLab Research Institute, Anglia Ruskin University, Cambridge, United Kingdom, [2]NowHere Media, Berlin, Germany, [3]School of Psychology, Sport and Sensory Sciences, Anglia Ruskin University, Cambridge, United Kingdom, [4]Centre for Psychological Medicine, Perdana University, Kuala Lumpur, Malaysia

Introduction: *Reviving Kusunda* is a project that explores the utility of immersive storytelling for the revitalisation of an endangered indigenous culture and language in Nepal called Kusunda. As part of this project, we compared the impact of a virtual reality (VR) experience (entitled *Kusunda: Speak to Awaken*) and a 2D companion film on audience engagement and emotional connections with the personal stories portrayed in both formats.

Methods: Data were collected through three focus groups conducted in the United Kingdom (n = 15) and three in Nepal (n = 15), as well completion of quantitative surveys by participants in the United Kingdom that were preceded by screenings of both formats (N = 100).

Results: The qualitative and quantitative results indicated the importance of integrating both VR and 2D film to maximize audience engagement. Specifically, while viewing the 2D film as a group appeared to promote group identity and wider access, the VR format was considered more effective for immersing oneself within the environment and developing an emotional connection, building on an understanding of how members of the Kusunda community live. Similarly, the quantitative results showed that participants found both formats to be highly informative and in terms of knowledge gain, although the VR format was more effective than the 2D format in terms entertainment and enjoyment, and emotional engagement.

Discussion: Overall, these results suggest that using a multi-pronged approach that leverages the benefits of both VR and 2D film can be effective in attempts to preserve endangered languages and heritage.

KEYWORDS

virtual reality, cultural heritage, endangered languages, 3D visualization, digital heritage, multimedia

1 Introduction

Endangered intangible heritage and language communities hold unique traditional knowledge that supports a distinct sense of place-making, identity, and long traditions of cultural and heritage sustainability (Braber and Howard, 2023; Orlove et al., 2022). As these communities are often marginalized and small in scale, their cultural heritage and its geographies are increasingly threatened by globalization and socio-economic development (Braber and Howard, 2023; Eaton and Turin, 2022; Smith, 2006). A key but often neglected issue in minority language and cultural heritage studies is the engagement of a variety of social actors—not only local, but supranational and global—who advance claims to, and participate in, the sustainable management of a given linguistic (Pellegrino, 2021) and cultural heritage (Labadi, 2022). Additionally, the internal fragmentation endemic to marginalized contexts and recurrent intergenerational tensions at play within them are often overlooked (Olko and Sallabank, 2018; O'Rourke and Pujolar, 2019).

Kusunda serves as a poignant example of an endangered heritage and language teetering on the brink of disappearance (van Driem, 2007). The Kusunda people constitute an indigenous ethnic group who predominantly lived a semi-nomadic lifestyle in the forests of central and western Nepal, and historically spoke a language that is commonly known by the same name. Despite possibly being the oldest and most indigenous people of Nepal, the Kusunda community have experienced a decades-long decline as a result of a range of factors, including inward migration by other ethnic groups, their displacement from historic villages, decreased livelihood opportunities, and imposition of the Hindu caste system and Nepalese national integration policies (Bodt and Aaley, 2024). This, in turn, has diminished the number of people who are able to speak Kusunda (Reinhardt, 1976), who are now estimated to be in the single digits. The near extinction of the Kusunda language is especially important given that it is a language isolate (Shafer, 1954; Watters, 2006) and has many features that make it unique within the wider Asian context (Bodt and Aaley, 2024; Watters, 2006).

Digital technologies provide a powerful opportunity to revive endangered intangible heritage by fostering increased participation and engagement from communities and other stakeholders in revitalisation efforts and practices (Edmonds et al., 2005; Galeazzi et al., 2022; Schnabel et al., 2016). However, to date, only a handful of attempts have been made to create prototype digital applications for the preservation of endangered heritage and languages. These attempts have predominantly focused on the development of online archives or platforms for the preservation of grammatical or lexical aspects of a language (INNET, 2014; Nordhoff, 2020) without exploring communities' engagement with their intangible heritage and language using innovative technologies, such as extended reality (XR) (Milgram et al., 1994; Rauschnabel, 2018). The challenges of high costs and poor distribution potential have further hampered the use of XR technologies in this space (Chuah, 2019).

Despite the low uptake, studies in the education and training sectors have indicated that virtual reality (VR) is often more effective than traditional learning methods in terms of emotional engagement (Dubovi, 2022; Flavián et al., 2021) and learning outcomes (Feng and Ng, 2023; Hua and Wang, 2023). For instance, some studies have shown that the use of VR is more effective than 2D film at improving knowledge acquisition, understanding of a topic, and the promotion of positive emotions (Allcoat and von Mühlenen, 2018). In part, this likely reflects the greater ability of VR to truly represent natural environments in terms of the scope and scale of settings, transport individuals away from their physical surroundings, and engage multiple sensory modalities (Hedblom et al., 2019). However, much of this research adopts a comparative approach—considering VR in relation to 2D film (Schutte and Stilinović, 2017; Ding et al., 2018; Herrera et al., 2018; Cohen et al., 2021)—which limits opportunities to consider the potential benefits of employing both approaches in a multi-pronged strategy to maximize audience engagement.

To begin addressing this research gaps, our study undertakes an analysis of viewers' feedback on emotional connection and understanding of an immersive VR experience (entitled *Kusunda: Speak to Awaken*: NowHere Media, 2020) and a 2D documentary film. The focus is on the heritage and language of the Kusunda community in Nepal, as portrayed through the eyes and testimonies of three members of the Kusunda community: Gyani Maiya Kusunda, Hima Kusunda, and Lil Bahadur Kusunda. Our investigation examines the degree to which VR and 2D film are able to promote emotional engagement with the theme of endangered indigenous languages and heritage.

To provide a comprehensive answer to this question, we adopted a mixed-methods strategy involving qualitative and quantitative analyses. In the first instance, participants were invited to view both the VR and 2D film formats before taking part in focus group discussions aimed at eliciting their views about the experience. In terms of the latter, we invited a new set of participants to experience one of the formats before completing a brief survey assessing their experiences. In broad outline, while we expected the VR format to promote a greater sense of immersion and emotional engagement, we also expected that both formats would promote knowledge gain. Qualitatively, we also considered the benefits of a multi-pronged approach that leverages the strengths of each format to maximize audience engagement and participation in the revitalisation efforts of endangered heritage and languages.

2 Materials and methods

2.1 Overview

Kusunda: Speak to Awaken is a VR experience that was created to document the Kusunda community's efforts to revitalize their language and to ensure that customs and traditions related to Kusunda heritage are preserved. The present study considers the impact of the VR experience and a 2D companion film which was created from footage used in the VR experience. The evaluation study was organized around two viewing experiences: one in VR, which included voice-based interaction, offering users an opportunity to learn and speak Kusunda words or phrases in order to move further in the narrative; the other, a 2D documentary film created subsequent to the VR experience for this evaluation, as a means of comparison. There were two methods of data collection: focus groups, with a discussion after viewing both films (i.e.,

FIGURE 1
Volumetric filming in front of a green screen. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

qualitative analysis), and individual viewings, where participants viewed either format or both before completing a brief survey (i.e., quantitative analysis).

2.1.1 Data collection and co-creation with Kusunda community

The interactive virtual reality experience *Kusunda: Speak to Awaken* was directed and produced by documentary makers Gayatri Parameswaran & Felix Gaedtke of NowHere Media with contributions from three members of the Kusunda community – Gyani Maiya Kusunda, Hima Kusunda, and Lil Bahadur Kusunda. The experience deployed various techniques, such as voice-based interactions, volumetric video, photogrammetry, and motion capture animations, all integrated within a game engine (Figure 1). Further information about the development of the experience can be found in Video 1 (Behind the scenes video describing the production of the VR experience).

2.1.2 Qualitative evaluation

In total, we conducted six focus group interviews (see Table 1). Three of the focus groups were conducted at Anglia Ruskin University (ARU) in the United Kingdom, with students in both undergraduate and postgraduate courses ($n = 15$). A further three focus groups were conducted in Nepal, close to the area where the Kusunda community live. Two of these were conducted with members of the Kusunda community now living in Lamahi ($n = 10$) and a further focus group ($n = 5$) was conducted with five representatives of the Language Commission Nepal in Kathmandu.[1]

The focus groups were either conducted offsite by industry partners within local community spaces (Nepal) or by researchers at the

university (UK). All focus group discussions were audio- and video-recorded. Each discussion with participants took between 45 min and 1 h.

During the focus group sessions, both formats were shown to participants before being discussed between participants and an interviewer, so as to determine the potential merits and drawbacks of both formats. The participants watched the 2D documentary film together as a group and the VR experience individually through a headset (Figure 2). After all participants viewed both versions, a trained interviewer led semi-structured discussions aimed at understanding viewing experiences. All focus group interviews were transcribed verbatim for analysis. For the Nepali focus group sessions, the transcripts were created from the video- and audio-recordings and were provided by a research assistant in the local region. The research assistant translated these transcripts into English and then provided these to the project leader.

The researcher then read all six focus group transcripts, examining participants' specific responses highlighting their emotional connection with the protagonists in the narrative or the wider subject matter of the VR versus 2D film itself. The researcher also considered any feedback regarding preferences for the VR or 2D film formats and what worked more effectively in one format in contrast to the other. The data was analyzed using thematic analysis as outlined by Kiger and Varpio (2020).

2.1.3 Quantitative evaluation

The study also included viewings by participants in the United Kingdom who were recruited opportunistically from campus locations ($N = 100$). The majority of participants (88%) were of white or British white ancestry. Having watched either the VR ($n = 50$) or the 2D film ($n = 50$), participants were asked to complete a survey assessing emotional engagement, interest in the topic and experience, effectiveness in promoting awareness of the importance of preserving endangered heritage and languages. Survey responses formed the basis of quantitative analyses comparing responses to the VR

---

1 http://languagecommission.gov.np/

presentation and the 2D film format. The analysis of responses was conducted using a multivariate analysis of variance (MANOVA) to examine whether the mean responses of the two groups (i.e., VR and 2D Film) significantly differed across survey items.

3 Results

3.1 Qualitative evaluation (focus groups)

The findings suggested that, while many participants expressed a preference for the VR experience in terms of engagement, emotional connection with the community, and generating interest in the topic, both formats induced different types of connection with viewers. Based on thematic content from the focus group sessions in the UK and Nepal, different facets of understanding of endangered heritage and culture were explored when the material was viewed individually (VR) or as part of a group experience (2D film). The analysis is divided under four themes: (1) "Feeling Kusunda": The viewer as a character in the film; (2) Interaction with the Kusunda language; (3) Immersion and 'physicality' of the VR experience; and (4). Narrative and content. We also distinguish between the focus groups conducted in the UK and those conducted in Nepal.

3.1.1 Focus groups in the UK

3.1.1.1 "Feeling Kusunda": the viewer as a character in the film

When experiencing the VR, participants who took part in focus groups in the UK noted how they felt like a character in the film. They reported feeling deeply immersed in the action, establishing a stronger emotional connection with both the story and the community portrayed. The VR format effectively facilitated a more profound understanding of the Kusunda lifestyle, such as living in the forest, as viewers were invited to interact with elements intrinsic to that way of life. According to them, the visual representation of the present (i.e., 3D realistic replicas of current landscapes) and the past (i.e., memories of Kusunda life in the jungle) were very effective in making them feel part of the Kusunda community (e.g., a scene where you are around the fire in the jungle with the community; Figure 3). Nonetheless, some participants noted how this immersion could distract from the main content of the film, namely, gaining insights into Kusunda culture and language.

TABLE 1 Focus groups' participants.

Focus group number	Location	Participants
FG1	Anglia Ruskin University	5 (MA in Sustainability)
FG2	Anglia Ruskin University	5 (BA in Digital Media)
FG3	Anglia Ruskin University	5 (BA in Artificial Intelligence)
FG4	Nepal (Kathmandu)	5 (Representatives of the Language Commission Nepal)
FG5	Nepal (Lamahi Region)	5 (Community members)
FG6	Nepal (Lamahi Region)	5 (Community members)

… when we watched it in VR it made me feel I am inside the very thing, inside the story, it made me feel I am progressing with the story. And the fact that there was an encounter with the bear, Hima recites some poem/song, the things about their day-to-day life, it felt like I was watching it from a very close perspective, the whole event is happening right in front of me, it made me feel I reached the very location, so I feel this technology can be more effective in today's time." (Participant 5 from FG4)

Participants observed that the use of animation to represent scenes from the past was notably more impactful in VR, especially in terms of employing color to convey emotion. A participant from FG1 expressed a desire to reach out and touch those elements within VR, making the process a far more tangible experience. Specific scenes, such as a communal gathering around the campfire, were highlighted by participants as particularly personal and effective at focusing the attention and providing an engaging display of oral histories. Additionally, there was a recurring theme of "feel[ing] the scene" (FG2), where figures in the film would make direct eye contact with the viewer, thereby increasing the sense of immersion.

The animated figures in the VR were deliberately stylised, and participants felt it appropriate that there was an element of abstraction to these images, rather than attempting to ensure these figures looked as lifelike as possible. One participant articulated that there was "something nice about being able to fill in the blanks a little bit" (Participant 1 from FG1). Participants recognized that overly realistic depictions of figures could hinder identification and engagement with the characters, diminishing the likelihood that they would engage and subsequently recall these elements of the VR after viewing. Conversely, participants deemed the use of real characters effective for visually representing the present, fostering an emotional connection with members of the communities, and illuminating their needs and concerns (Figure 4). This approach heightened the emotional resonance for participants, enhancing their overall engagement with and recall of the VR experience.

So, when the girl says that she wants us to learn the language to preserve it from disappearing, you feel like you're not sure

FIGURE 2
Participant experiencing the application through a VR headset.
© NowHere Media.

FIGURE 3
3D animation representing members of the Kusunda community around the fire. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

FIGURE 4
Volumetric capture of Hima and Lil Bahadur overlapped to the 3D reproduced scene. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

you have to, but you feel her. You feel what she wants because she's with you. (Participant 5 from FG3)

The significance of sound for immersion was underscored by participants, emphasizing its importance on par with visuals. For instance, participants conveyed that a song in the Kusunda language resonated more powerfully within the VR. The heightened sense of connection was attributed to a perceived personal connection with the singer, emphasizing the effectiveness of the VR medium in fostering a more intimate and meaningful engagement with the cultural expression. Members of FG2 also specifically highlighted how the VR excelled in capturing intangible aspects of heritage, emphasizing the significance of the soundscapes, including the distinct sound of the Kusunda language and sound of the environment in which the protagonists lived.

But then like feeling within the VR that you're so immersed in the sound which is like feels, like, I don't know if this is quite what you're getting at. But for me it's like I felt like the sound world and even things like when the granddaughter was speaking and you had like a radio on in the background, is that kind of immersive. (Participant 1 from FG1)

3.1.1.2 Interaction with the Kusunda language

The VR experience also provided an opportunity for direct interaction with the Kusunda language itself, a component that participants found a particularly powerful part of the experience. Some of them considered the VR medium more effective than the 2D film in terms of language learning and word retention at the conclusion of the experience: "What I noticed differently at the end… I could remember some of the words, like the word for forget… When

I watched the (2D) video I did not even realize how these words were at the end" (Participant 4 from FG3).

During the VR presentation, viewers were prompted to repeat specific Kusunda words to activate various elements of the experience. However, in the event of technical issues with this interactivity, the VR would proceed even without receiving the required prompt. This interactive feature was absent in the 2D film version, as though words with the phonetic notation were included, the viewer is not invited to repeat these words aloud.

Participants expressed a sense of inherent power in being able to speak words from the Kusunda language. The interaction with the language was considered significant by both groups, with a preference for using voice commands over hand controls. While some participants felt that this interactivity did not necessarily deepen emotional involvement (FG1), the majority of them recognized that speaking words from the language fostered a heightened emotional understanding of the life of one of the central figures within the documentary, who had lost his language and, consequently, his heritage.

> So, if you're living that lifestyle, and then as soon as you are taken out of that environment, you have to learn those words to describe stuff. So, you forget your own language fairly quickly. (Participant 1 from FG2)

This observation underscores the emotional connection which helps viewers identify a personal outcome for the issues concerning endangered languages. Participants expressed that through the interactive language component, they could actively practice and gain a deeper understanding of pronunciation practices. Importantly, they reflected on the potential benefits of such an experience for the younger generations, envisioning it as a valuable tool for learning and preserving the endangered language.

> …even when the interaction part was done, I was still thinking how beneficial it could be if the younger aged people could watch it. (Participant 4 from FG3)

Participants also noted that saying the words aloud during the VR experience was seen as an active contribution to keep this language alive. Rather than just relying on written resources, viewers and future audience retain greater information as regards pronunciation and tone when a recorded version is created. As such, participants noted how the VR provides an isolated environment and a personal experience with the language, thereby making it more effective at preserving it.

3.1.1.3 Immersion and 'physicality' of the VR experience

The benefits of using VR were associated with its 'physicality', offering a unique perspective on the narrative. A participant from FG1 noted the distinctiveness of VR in emphasizing the importance of the language and bringing aspects of the recent past to life through immersion. Interestingly, many viewers perceived the VR experience as having a shorter running time than the 2D film, a perception attributed to the interactivity, as expressed by a participant 4 from FG2: "It actually seemed shorter than watching it on the screen." The interactive elements in VR contributed to a sense of time passing swiftly.

However, there were also comments about the drawbacks of wearing of the VR headset. Some participants found the weight of the headset to be distracting and even discomforting over time: "I found that the VR headset was quite distracting because it just kind of starts pushing on my face after a while and it gets quite painful" (Participant 2 from FG1). Those participants wearing glasses reported additional issues with the weight of the headset pushing on the face, which caused some level of discomfort. One participant in FG2 also reported issues with eyestrain while another mentioned that they were concerned while moving due to the cable attaching the headset to the laptop. Due to technological limitations, it was not possible to have free navigation, though participants in FG2 were unsure whether that would be necessary for a VR experience of this nature.

These responses suggest a delicate balance between the immersive physicality of VR, as seen in previous sections, and the potential for physical discomfort or distraction from the narrative content. One participant in FG2 reported becoming more aware of his body and surroundings after 10 min of wearing the headset—such as how he was sitting or where to put his hands—leading to a slight loss of focus on the content. Additionally, participant 1 in FG1 highlighted that for those with limited experience in VR, the technological element itself could be distracting: "And if you have not had much experience in VR, then, again, it's not necessarily the content, but it's just the technological element that distracts a bit from the emotion and the meaning."

3.1.1.4 Narrative and content

Participants also identified that a balancing act occurred between the ability to interact with the VR environment and the drawbacks associated with wearing the headset, which, at times, caused some discomfort and distraction. The isolation provided by the VR headset was perceived differently by participants: (a) Some participants recognized the benefit of the 2D film, emphasizing the possibility of watching it in a group. This communal viewing was seen as advantageous for discussions regarding the documentary, creating a shared event, and enhancing the audience's understanding of the narrative themes, particularly in terms of group identity; (b) Conversely, other participants considered the isolation imposed by the VR headset less distractive compared to the group experience occurring with the 2D film.

> VR provides… more opportunity to connect emotionally to the subject and the people. When I watched the video in 2D instead… I started to lose focus on the storyline and on the people… I turned right, and I saw other participants who were watching how they felt the same way… when I watched the VR, I didn't have this feeling. (Participant 3 from FG3)

In general, VR elicited a greater sense of awareness and emotional engagement with the protagonists and their attempts to preserve the language and culture of the Kusunda community. On the other hand, the 2D film facilitated an easier experience with the contents, serving as an important document for preserving a heritage that is now under threat. The 2D film, in particular, offered insights into how the Kusunda culture interacts with the modern world, as demonstrated through the interaction between Lil Bahadur and Hima with the film crew (Figure 1).

The possibility of using both formats is highlighted by one participant in FG2 who suggests watching the 2D film presentation before the VR to provide a wider contextual framework before being

immersed in the VR environment. Additionally, given that VR is still a novel experience for some viewers, participants in FG2 saw it as an intriguing hook for those with only a passing interest in the subject depicted in the film. Interestingly, all participants in FG2 expressed a preference for VR over the 2D film, considering it the more absorbing and engaging of the two formats, especially for individuals new to VR technology.

What participants found particularly effective in the 2D film was the personal account at the heart of the narrative. However, the animation was singled out as more effective and engaging in the VR format compared to the 2D film. The immersive nature of VR seemingly heightened the impact of the animated elements, providing a more compelling and absorbing experience for the participants. This observation underscores the unique capabilities of VR in enhancing the effectiveness of certain visual elements, such as animation, contributing to a more immersive and captivating storytelling experience.

> I found the animated, little, kind of like, seeing the animated scenes were a lot better on VR, particularly the one with the bird. Yeah, because they were like right in front of you and you could watch it rather than being slightly further out. (Participant 2 from FG2)

The animated features in the VR format were noted to be distracting from the filmed interviews with people, according to participants in FG1. However, participants in FG2 expressed the possibility that changes in how material was filmed or presented in VR could mitigate this distraction and potentially enable a more personal connection.

In terms of language presentation, the 2D film used subtitles, while the VR relied on dubbing, resulting in a mixed response. Participants in FG2 found subtitles distracting and occasionally disconnected from the spoken words of the interviewees. Conversely, participants in FG1 preferred subtitles, as they allowed for clear hearing of the interviewee's

voice. In the VR, participant 3 (FG1) noted, "I've got the English voice in my ear, rather than just hearing his voice" suggesting that hearing the interviewees directly established stronger emotional connection. However, it is worth noting that this response might vary among individuals, and further research is needed to determine the prevailing sentiment on this matter.

Participants felt the presentation of the 2D film prompted viewers to contemplate broader questions about group identity and its significance in relation to language. Participants emphasized that responses to group identity would differ across various locations, with some groups being more isolated geographically and culturally than others. As participant 1 in FG1 articulated, "this was a very small language in a very small community whereas other languages might not be quite the same."

3.1.2 Kusunda participants

3.1.2.1 "Feeling Kusunda": the viewer as a character in the film

Many of the points raised by the UK focus groups found resonance among the Nepali focus groups, with additional insights reflecting the broader concerns of those with pre-existing knowledge of the Kusunda community. In VR, participants noted that the entire forest was visible, in contrast with the limited view in the 2D film, enhancing their connection with the environment and the experience of speaking the Kusunda language.

Participants in Nepali FG4 described the VR experience as akin to "feeling Kusunda" expressing a notable emotional attachment to the immersive environment. Participants in the Nepali focus groups, in general, reported a greater emotional connection when viewing the VR, potentially attributed to their prior knowledge of the Kusunda community and the interviewees within the VR itself. For example, one viewer in FG5, who was a family member of Gyani Maiya Kusunda, felt a profound connection as if she was in front of her again (Figure 5).

FIGURE 5
Interview of Gyani Maiya Kusunda in the 3D VR experience. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

…I felt I was talking to her (Gyani Maiya), she has passed away now, but when I watch her in VR, I felt like I was talking to her. I felt like I was sitting with her. I felt like she was close, I was tearing up, but I could not cry. When she was alive, she did lots of things for the language, when she was alive the language was alive with her. After her the language took its downfall. (Participant 2 from FG5)

This participant's personal connection with the narrative, her ability to speak some Kusunda, and her relationship with Gyani Maiya contributed to the heightened emotional engagement facilitated by the VR. Community members also echoed comments by participants in focus groups in the UK on the effectiveness of the graphic representations of memories from the past, depicting Kusunda life in the jungle. They mentioned how these representations allowed them to see for the first-time stories they had previously only heard from their grandparents.

The events shown here were similar to the stories I used to hear… the way they used to live, the way they used to hunt, the kind of clothes they used to wear, I could relate to that…when I watched this today, I felt like I was watching all the stories grandmother used to tell me, they were in front of my eyes, as if they were real. (Participant 2 from FG5)

3.1.2.2 Interaction with the Kusunda language

Participants from the Kusunda community pointed out the potential and benefits of creating a similar VR experience for other endangered heritage and languages in Nepal. However, they also acknowledged the importance of using both formats (VR and 2D film) to address accessibility issues, considering factors such as access to technology and familiarities with it.

…even if VR is not possible to show to everyone, you could show the TV documentary (i.e., 2D Film) to as many people as you can,

also some old people might get scared of all this… the equipment is also not accessible and they are expensive, but if we can show this in TV everyone could watch in their home, old generation people from the community can watch comfortably… it would be better if more people could watch it. (Participant 3 from FG6)

Participants identified the importance of preserving native languages as a fundamental human right. They recognized the potential of the VR medium as a significant asset for the preservation of their language and culture. As the VR experience involves viewers pronouncing Kusunda words, participants believed that this interactivity played a central role in continuing the legacy of Kusunda and ensuring the broader preservation of the language. The interactive aspect of the VR experience was seen as a key factor in engaging a wider audience and fostering a more active participation in language preservation efforts.

…about the topic of language, it has given huge importance to language, if we don't speak, if we don't pronounce the Kusunda words it doesn't move forward, and even in the middle and throughout the whole piece, it has moments where we need to speak Kusunda words, I liked it very much. Because of that, even a person who has never spoken this language before, he will have to speak the language, so by this technology, at least he is made to speak 5-6 words in Kusunda, that is a huge achievement. (Participant 5 from FG6)

Participants noted that placing the Kusunda language in the context of its history, as depicted in the VR experience, had the potential to bring the language back to life for modern audiences. Scenes such as the hunting (Figure 6), descriptions of customs, details about Kusunda diet, and information about traditional attire were mentioned as effective elements. Participants believed that the VR format, by contextualizing the language within its original settings, helped to identify important linguistic elements.

FIGURE 6
3D animation representing a hunting scene with members of the Kusunda community. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

Moreover, they suggested that the use of language within the VR narrative contributed a broader understanding of how endangered languages could be preserved. They expressed the belief that the VR experience could serve as a model for other endangered languages, especially those that are less known. Members of the Language Commission Nepal (FG4) echoed the sentiment that the language being at the heart of the VR documentary was one of the most powerful aspects and provides a very effective way to learn and retain Kusunda words.

> The most powerful thing of this experience was, I have learnt at least 7–8 Kusunda words, just by watching this. That is struck in my mind, I may not pronounce it correctly, but it has registered in my mind, like 'Maae' (Sangam) 'Mayak'! (Sanjog) like 'Gilang' and it is also similar to some of Magar words, that was also the reason it had an impact. (Participant 3 from FG4)

3.1.2.3 Immersion and 'physicality' of the VR experience

Participants in FG4 share a similar observation with those in FG2 and FG3 that the VR presentation seemed to have a shorter perceived duration compared to 2D film. This was attributed to the interactive nature of the VR experience, allowing participants to engage more actively with the virtual space. Additionally, participants noted specific instances, such as the rain scene, where the immersive quality of VR made them feel deeply connected to the environment, with one participant expressing, "…what I felt was, in the moment, in any moment if it was raining there, I would have felt I am getting wet in the rain myself." (Participant 4 from FG3).

FG5 participants highlighted the significance of such immersive experiences facilitated VR and the isolating effect it has on the viewer. They also found that the description and visuals used in the storytelling were more engaging in VR, providing a clearer connection between the past and the present. This connection was particularly important in the visualization of the connection to nature and to the jungle, with the animals having greater presence in VR (Figure 7).

However, akin to other groups, there were also concerns about the physicality of using VR and how immersion could ultimately distract from the important messages contained within the narrative. There were similar complaints regarding the use of the VR headset and that audiences may find this uncomfortable with participant 2 from FG5 noting how it felt almost "suffocate[ing]."

3.1.2.4 Narrative and content

Nepali participants, similar to the UK focus groups, expressed a preference for viewing both the VR and 2D film, believing it contributed to a better understanding of issues of central importance in the narrative and offer a more comprehensive experience as a result. However, in contrast to the UK focus groups, participants in FG6 felt that more aspects of the Kusunda culture could have been evident in the 2D film, though acknowledged the challenge of including extensive detail given the loss of much cultural information.

Participants in FG4 expressed less engagement with scenes featuring behind-the-scenes footage. Despite recognizing the importance of these scenes for storytelling, they still expressed an expectation that these moments should have been edited out: "…there is another part, while we were watching at one point the director says 'we have to take a break!' was that part really necessary?" (Participant 1 from FG4). This observation provides an interesting contrast with responses from FG1, who identified these scenes as important for contextualizing the personal story of Lil Bahadur and Hima and their environment. It could be argued that it is potentially less important to have these scenes for those living closer to the community, suggesting that those with greater awareness of the conditions may find these scenes less crucial.

3.2 Quantitative evaluation

As with the qualitative analyses in general, the quantitative analysis of the survey responses broadly showed that the VR presentation was more effective compared to the 2D film (total $N = 100$; Gender identity: 44 women, 55 men, 1 identified in another way; age $M = 31.88$, $SD = 10.99$). A MANOVA showed a significant omnibus effect of presentation format (VR vs. 2D), $F_{(8, 91)} = 2.65$,

FIGURE 7
3D animation representing Lil Bahadur with a bear. © NowHere Media, ZHdK, Poke Poke Collective, Fasad, INVR.

TABLE 2 Summary of the univariate ANOVA results.

Item	VR (n = 49)		2D (n = 50)		F	p	η_p^2
	M	SD	M	SD			
I found the VR/2D film entertaining and enjoyable	4.70	0.54	4.16	0.87	13.95	<0.001	0.13
I was emotionally engaged with the VR/2D film	4.48	0.65	4.10	0.86	6.20	0.014	0.06
I learned something new from watching the VR/2D Film	4.70	0.51	4.46	0.84	3.01	0.086	0.03
I would recommend the VR/2D film to my family and friends	4.62	0.75	3.98	0.98	13.42	<0.001	0.12
I found the VR/2D film very informative	4.50	0.61	4.36	0.80	0.96	0.330	<0.01
I found the VR/2D film thought-provoking and interesting	4.68	0.55	4.40	0.81	4.10	0.046	0.04
Watching the VR/2D film has made me want to find out more about endangered indigenous heritage and languages	4.30	0.79	3.96	0.85	4.26	0.042	0.04
I believe the VR/2D film is effective at promoting awareness of the importance of preserving endangered heritage and languages	4.58	0.64	4.20	0.90	5.87	0.017	0.06

$p = 0.012$, $\eta_p^2 = 0.19$, and the univariate ANOVA results are reported in Table 2.

The univariate results showed that participants (a) found the VR presentation more entertaining and enjoyable (large effect); (b) were emotionally engaged with the VR presentation (medium effect); (c) were more likely to recommend the VR presentation to family and friends (large effect); (d) found the VR presentation more thought-provoking and interesting (small effect); (e) were more likely to want to find out more about endangered indigenous heritage and languages after the VR presentation (small effect), and (f) believed the VR presentation to be more effective at promoting the importance of preserving endangered heritage and languages (medium effect). Conversely, there was no significant difference between the VR and 2D presentation formats in terms of learning something new and in terms of perceptions of informativeness.

Mean responses were generally high across all items and for both formats (Ms for VR ranged from 4.30 to 4.70 and for 2D format from 3.96 to 4.46). Overall, these results suggest that participants found both formats to be highly informative and in terms of knowledge gain. However, the VR format was more effective than the 2D format, particularly in terms entertainment and enjoyment, and emotional engagement. The VR format was also slightly more effective than the 2D format in terms of provoking interest, stimulating further interest, and in terms of perceptions of effectiveness. Participants were also much more likely to recommend the VR version to their friends and family.

Separately, we also added gender to the analyses, but there was no significant main effect of gender, $F(16, 178) = 1.34$, $p = 0.177$, $\eta_p^2 = 0.11$, and no significant gender by format type interaction, $F(8, 88) = 0.76$, $p = 0.641$, $\eta_p^2 = 0.06$. We also examined bivariate correlations between age and survey responses for each presentation format separately. In the VR format, older participants were more likely to have recommended the film to others ($r = 0.29$, $p = 0.039$) and to have found the film more informative ($r = 0.28$, $p = 0.046$). All other correlations were not significant. In the 2D format, none of the correlations reached significance. These results broadly suggest that participant demographics (gender identity and age) did not impact the main findings.

4 Discussion

Digital technologies offer powerful opportunities to revive, revitalize, and preserve intangible heritage (Edmonds et al., 2005; Galeazzi et al., 2022; Schnabel et al., 2016), but their deployment remains relatively piecemeal. To fill this gap, we considered how two distinct formats—VR and 2D film—engage different audiences in terms of the revitalisation of an endangered heritage and language, namely the Kusunda of Nepal. In broad outline, our qualitative results indicated that participants generally preferred the VR format, but there was also evidence that the 2D format could be effectively leveraged to promote understanding and knowledge gain. Likewise, our quantitative results showed that the VR format was more effective than the 2D format in terms of emotional engagement, but both formats were equally effective in terms of knowledge gain. Based on these results, we suggest that a multi-pronged approach that leverages the benefits of both VR and 2D film may be beneficial for reviving and revitalizing intangible heritage and language.

The results of our qualitative analyses showed that participants broadly preferred the VR presentation over the 2D format, highlighting the importance of interactivity and immersion in engaging viewers in the subject matter. This was true in relation to both the focus groups that were conducted in the UK, as well as those held in Nepal., Likewise, our quantitative data from individual viewings consistently indicated that the VR format was more effective than the 2D format at invoking interest, creating a connection with the viewer regarding the topics discussed, and inspiring audiences to pursue an interest in endangered languages and heritage after watching the presentation.

These results are broadly consistent with earlier work suggesting that VR may be more effective than 2D film at promoting emotional engagement with content (Schutte and Stilinović, 2017; Ding et al., 2018; Herrera et al., 2018; Cohen et al., 2021). This may reflect the fact that VR offers greater vividness and interactivity, including in terms of the extent and realism of actions that can be performed in the virtual environment (Breves and Heber, 2020; Slater and Sanchez-Vives, 2016). In terms of our VR experience specifically, the ability to interact and thereby control the flow of the narrative may also have enhanced feelings of immersion, which in turn foster feelings of social presence or a sense of "being there" and experiencing what was

simulated as real (Wirth et al., 2007). In short, our results suggest that the VR technologies can be effectively deployed in efforts to engage audiences with the revitalisation of endangered heritage and language.

However, both our qualitative and quantitative results also showed that there were benefits to the use of the 2D format. For instance, participants in our focus groups recognized VR as a novelty for many audiences and further cautioned about the need to balance immersion with the amount of discomfort or distraction caused by wearing the headset. Moreover, with the 2D film, despite offering less immersion in the viewing experience, 'physicality' was not a distraction from the personal stories and content of the documentary, as noted in discussion about the VR experience. In a similar vein, our quantitative results showed no significant differences between the VR and 2D formats in terms of knowledge gain and perceived informativeness. In short, our findings suggest that there are real benefits to the 2D format, which in some cases may be comparable to the VR format.

Based on these results, it may be suggested that both VR and 2D film could play important roles in terms of the preservation of endangered intangible cultural heritage or language. Of course, one of the challenges now is determine which specific format may be most efficient in specific contexts in terms of preserving aspects of intangible cultural heritage. However, one avenue emerged from our focus groups, which emphasized the importance of appropriate viewing environments to minimize distraction. As such, it may be that the VR format could be deployed in ideal circumstances and where adequate resources exist, whereas the 2D format may be more appropriate where resources are limited or where VR technologies are not available. Indeed, focus groups with our Nepalese participants highlighted the need to disseminate the film widely, which would be facilitated more easily through use of the 2D format.

In reality, of course, efforts at revitalizing endangered cultural heritage and language may best be served by multi-pronged initiatives that include multiple modalities and technologies (Feng, 2024). Indeed, our focus group participants noted how important such multi-modal projects are, not just in creating a final product, but in revitalizing and securing a group identity through co-creation and ensuring members of a community are part of the creation process at every stage to ensure that experiences remain authentic and accurate. In planning future projects, it will also be important to play to the strengths of specific technological formats in terms of emotional engagement. In our case, viewing the 2D film as a group helped with connection to group identity but the VR was more effective for immersing the viewers within the environment, so helped develop an emotional connection, building on an understanding of how individual members of that community live day-to-day. As a result, participants noted how crucial it was to 'feel' like the Kusunda, opening up a wide range of possibilities of how VR is an important tool for audience engagement.

4.1 Limitations and future directions

This study provides an overview of the emotional impact of the VR experience compared to a more traditional 2D format and serves as a starting point in moving forward when considering how to manage engagement with endangered heritage and language communities on this important topic. Nevertheless, there are several limitations of the present work that should be considered and addressed in future research. First, although the project provides a timely and much-needed comparison of VR and 2D film, the comparison of only two approaches to preserving the Kusunda heritage and language represents a limitation of the project, offering an incomplete perspective. This limitation arises mainly due to the complexity of developing the different digital creative outputs and the resources available for this project. Therefore, in the future, it will be important to scale up this project to compare and co-create multiple media formats.

In a similar vein, our findings also raise more complex questions for exploration in future studies. For instance, assuming co-creation with communities from the beginning of the process is an unavoidable element, then how can XR technologies be effectively used within a community-led participatory framework to empower communities and effectively address issues of inequality? Additionally, given participants' observations about the potential isolation of VR compared to the communal experience of watching a film, might the use of immersive technologies that combine the 'virtual' with the 'physical' worlds have an impact on people emotional engagement and interaction with endangered heritage and languages? Ultimately, our results show how the use of multiple formats could play the strengths of each presentation and increase access to information and contents. Also, other formats, such as augmented reality, may prove beneficial in raising awareness about endangered heritage and languages. By combining the 'virtual' and 'physical' worlds, these formats could find applications in museum exhibits, allowing for interactive experiences with artifacts or digitally produced reproductions in communal spaces. Such an approach might mitigate some of the challenges associated with VR headsets. Although participants universally acknowledged how the VR medium was able to vividly represent and bring intangible aspects of heritage to life. As noted by a participant in FG3, VR enhances access and engagement, effectively "making the intangible [heritage], tangible."

In addition, the present work was concerned only with qualitative and quantitative results drawn solely from a very small subset of the communities in Nepal and the UK, which may not provide a comprehensive account of viewers' experiences. One way to extend this in future work would be by collecting additional data from different cultural context and other endangered heritage and languages, as well as a higher number of participants that may provide a fuller account of viewers' experiences during the immersive experience. In a similar vein, it will be important to assess the extent to which engagement with VR and other media formats is able to alter attitudinal, cognitive, and behavioral outcomes, a task suited to large scale quantitative research methods. Finally, it should be noted that the present work was limited to qualitative and quantitative experiences solicited immediately after viewing the two media (VR and 2D Film). In future work, it may be useful to assess longer-term outcomes as a result of viewing the immersive film and other media formats.

5 Conclusion

Despite these limitations, the present study shows that both VR and 2D film can be effectively used to promote knowledge gain and perceived informativeness in terms of endangered cultural heritage and language. However, VR was considered more effective than 2D film in fostering a deeper sense of connection with the Kusunda communities

and their stories, enabling participants to empathize and "feel" like the Kusunda people. This highlights the potential for integrating VR with other media in revitalisation efforts aimed at creating closer emotional interactions with endangered heritage and language communities.

While the use of *Reviving Kusunda* as a tool in and of itself needs further evaluation across diverse populations and through comparison with multiple media formats, as outlined above, the overall framework here can serve as a valuable guide for researchers and practitioners aiming to create deep emotional connection with communities, leveraging emotional resonance and specific needs and concerns. Specifically, the use of a co-creation framework is particularly vital, as it ensures that the voices, experiences, and understandings of individuals are meaningfully embedded in any intervention. This is crucial to ensure those who are part of minorities and marginalized communities are not excluded. By doing so, it enhances the impact of multi-pronged approaches that leverage the benefits of integrating and selecting multiple media formats tailored to the specific needs and concerns of each endangered heritage and language community.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Faculty of Arts, Humanities, Education and Social Sciences (AHESS) Research Ethics Panel, Anglia Ruskin University. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

FaG: Writing – original draft, Writing – review & editing. GP: Writing – original draft, Writing – review & editing. FeG: Writing – original draft, Writing – review & editing. VS: Writing – original draft, Writing – review & editing.

References

Allcoat, D., and von Mühlenen, A. (2018). Learning in virtual reality: effects on performance, emotion and engagement. *Res. Learn Tech.* 26:2140. doi: 10.25304/rlt.v26.2140

Bodt, T. A., and Aaley, U. R. (2024). Decline of the Kusunda of Nepal. *Asian Ethno* 83, 301–332.

Braber, N., and Howard, V. (2023). Safeguarding language as intangible cultural heritage. *Int. J. Intangible Heritage* 18, 146–158.

Breves, P., and Heber, V. (2020). Into the wild: the effects of 360° immersive nature videos on feelings of commitment to the environment. *Environ. Comm.* 14, 332–346. doi: 10.1080/17524032.2019.1665566

Chuah, S. H. W. (2019). Wearable XR-technology: literature review, conceptual framework and future research directions. *Int. J. Technol.* 13, 205–259. doi: 10.1504/IJTMKT.2019.104586

Funding

Acknowledgments

Conflict of interest

GP and FG were employed at NowHere Media.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Cohen, D., Landau, D. H., Friedman, D., Hasler, B. S., Levit-Binnun, N., and Golland, Y. (2021). Exposure to social suffering in virtual reality boosts compassion and facial synchrony. *Comp. Hum. Behav.* 122:106781. doi: 10.1016/j.chb.2021.106781

Ding, N., Zhou, W., and Fung, A. Y. H. (2018). Emotional effect of cinematic VR compared with traditional 2D film. *Telemat. Info.* 35, 1572–1579. doi: 10.1016/j.tele.2018.04.003

Dubovi, I. (2022). Cognitive and emotional engagement while learning with VR: the perspective of multimodal methodology. *Comp. Edu.* 183:104495. doi: 10.1016/j.compedu.2022.104495

Eaton, J., and Turin, M. (2022). Heritage languages and language as heritage: the language of heritage in Canada and beyond. *Int. J. Heritage Stud.* 28, 787–802. doi: 10.1080/13527258.2022.2077805

Edmonds, E. A., Weakley, A., Candy, L., Fell, M., Knott, R., and Pauletto, S. (2005). The studio as laboratory: combining creative practice and digital technology. *Int. J. Hum. Comput. Stud.* 63, 452–481. doi: 10.1016/j.ijhcs.2005.04.012

Feng, H. (2024). The application of digital technology in the protection and inheritance of intangible cultural heritage. *Front. Psychol*, 11, 1–13. doi: 10.17501/23572744.2024.11101

Feng, B., and Ng, L. (2023). Facilitating writing performance of EFL learners via virtual reality: Immersion, presence, embodiment. Frontiers in Psychology, 14. doi: 10.3389/fpsyg.2023.1134242

Flavián, C., Ibáñez-Sánchez, S., and Orús, C. (2021). Impacts of technological embodiment through virtual reality on potential guests' emotions and engagement. *J. Hosp. Marketing Manage.* 30, 1–20. doi: 10.1080/19368623.2020.1770146

Galeazzi, F, Di Giuseppantonio Di Franco, P., Toulson, R., Camporesi, C., and Patel, S. (2022). Earthquakes, communities and heritage: telling stories of resilience through co-designed immersive media. *Vis. Stud.* 38, 101–118. doi: 10.1080/1472586X.2022.2102539

Hedblom, M., Gunnarsson, B., Iravani, B., Knez, I., Schaefer, M., Thorsson, P., et al. (2019). Reduction of physiological stress by urban green space in a multisensory virtual experiment. *Sci. Rep.* 9:10113. doi: 10.1038/s41598-019-46099-7

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., and Zaki, J. (2018). Building long-term empathy: a large-scale comparison of traditional and virtual reality perspective-taking. *PLoS One* 13:e0204494. doi: 10.1371/journal.pone.0204494

Hua, C., and Wang, J. (2023). Virtual reality-assisted language learning: a follow-up review (2018-2022). *Front. Psychol.* 14:1153642. doi: 10.3389/fpsyg.2023.1153642

INNET (2014). Innovative networking in infrastructure for endangered languages. Available at: https://cordis.europa.eu/project/id/284415/reporting (Accessed August 26, 2024).

Kiger, M. E., and Varpio, L. (2020). Thematic analysis of qualitative data: AMEE guide no. 131. Med. *Teacher* 42, 846–854. doi: 10.1080/0142159X.2020.1755030

Labadi, S. (2022). Rethinking heritage for sustainable development. London, UK: UCL Press.

Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. (1994). Augmented reality: a class of displays on the reality-virtuality continuum. *Telemanipulator Telepresence Tech.* 2351, 282–292. doi: 10.1117/12.197321

Nordhoff, S. (2020). "From the attic to the cloud: mobilisation of endangered language resources with linked data" in Proc. of the workshop about language resources for the SSH cloud. eds. D. Broeder, M. Eskevich and M. Monachini (Marseille: European Language Resources Asscoiation), 10–18.

NowHere Media (2020). Kusunda: speak to Awaken. Available at: https://www.speaktoawaken.nowheremedia.net/kusunda (Accessed August 26, 2024).

O'Rourke, B., and Pujolar, J. (2019). From new speaker to speaker: outcomes, reflections and policy recommendations from COST action IS1306 on new speakers in a multilingual Europe. Cardiff: IAITH.

Olko, J., and Sallabank, J. (2018). Revitalising endangered languages: a practical guide. Cambridge, UK: Cambridge University Press.

Orlove, B., Dawson, N., Sherpa, P., Adelekan, I., Alangui, W., Carmona, R., et al. (2022). "ICSM CHC white paper I: intangible cultural heritage, diverse knowledge systems and climate change" in Contribution of knowledge systems group I to the international co-sponsored meeting on culture, heritage and climate change. Discussion paper. ICOMOS & ISCM CHC (Charenton-le-Pont, France & Paris, France).

Pellegrino, M. (2021). Greek language, Italian landscape: Griko and the re-storying of a linguistic minority. Hellenic Studies Series 89. Washington, DC: Center for Hellenic Studies.

Rauschnabel, P. A. (2018). Virtually enhancing the real world with holograms: an exploration of expected gratifications of using augmented reality smart glasses. *Psychol. Market.* 35, 557–572. doi: 10.1002/mar.21106

Reinhardt, J. (1976). The ban rajas – a vanishing Himalayan tribe. *Contributions to Nepalese Stud.* 4, 1–21.

Schnabel, M.A., Aydin, S., Moleta, T., Pierini, D., and Dorta, T. (2016). Unmediated cultural heritage via Hyve-3D: collecting individual and collective narratives with 3D sketching. Paper presented at the 21st international conference of the Association for Computer-Aided Architectural Design Research in Asia (CAADRIA), Melbourne, Australia, March 30–April 2, 2016.

Schutte, N. S., and Stilinović, E. J. (2017). Facilitating empathy through virtual reality. *Motiv. Emotion* 41, 708–712. doi: 10.1007/s11031-017-9641-7

Shafer, R. (1954). Ethnography of India. Wiesbaden: Otto Harrasowitz.

Smith, L. (2006). Uses of Heritage. London: Routledge.

Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Front. Robotics AI.* 3:74. doi: 10.4324/9780203602263

van Driem, G. (2007). "Endangered languages of South Asia" in Handbook of endangered languages. ed. M. Brenzinger (Berlin: Mouton de Gruyter), 303–341.

Watters, D. (2006). Notes on Kusunda grammar: a language isolate of Nepal (Himalayan linguistics, archive 3). Kathmandu, Nepal: National Foundation for the Development of Indigenous Nationalities.

Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., et al. (2007). A process model of the formation of spatial presence experiences. *Media Psychol.* 9, 493–525. doi: 10.1080/15213260701283079