

Understanding 3D Data Videos: From Screens to Virtual Reality

Leni Yang*

Hong Kong University of Science and Technology

Wai Tong‡

Hong Kong University of Science and Technology

Zheng Wei¶

Hong Kong University of Science and Technology (Guangzhou)

Aoyu Wu†

Hong Kong University of Science and Technology

Xian Xu§

Hong Kong University of Science and Technology (Guangzhou)

Huamin Qu||

Hong Kong University of Science and Technology

ABSTRACT

Data storytelling explores how to communicate data insights to the general public engagingly and effectively. It combines the power of data visualizations and storytelling techniques and is popular in various media such as newspapers, interactive websites, and videos. Recently, virtual reality has brought new opportunities to enhance data storytelling with an incomparable sense of immersion. However, there exists a limited understanding of data stories in virtual reality (VR) as they are still in the early stage. In this paper, we investigated the idea of VR data videos by drawing inspiration from popular 3D data videos and studying how to transfer them from screens to VR. We systematically analyzed 100 highly-watched 3D data videos from Youtube and Tiktok channels to derive their design space. We then conducted a user study with 12 participants to explore the effects of four design factors on user experience, including varying camera angles, showing chart overview, animation, and using anchors. Specifically, participants watched 3D data videos in desktop and VR environments. We collected and analyzed their quantitative and qualitative feedback regarding the story’s understandability, memorability, engagement, and emotional effects. Results suggested that data videos in VR were significantly more appreciated than on desktops. We concluded with design implications for future applications and research on VR data videos.

Keywords: Narrative Visualization, Data Video, Virtual Reality

Index Terms: Human-centered computing—Visualization—Empirical studies;

1 INTRODUCTION

Data storytelling - a technique for telling stories with data graphics [50] - has been increasingly popular on the Internet. They often incorporate creative, elegant, and professional visual designs to attract and engage viewers. Practitioners and researchers have studied different genres of data stories, such as interactive articles, infographics, and videos [57]. With the advance in media technologies, the media for data storytelling expands from screens to immersive environments [29], which are thought to bring the engagement and immersion of the audience to a new level. Yet, we have limited empirical knowledge about data storytelling in immersive environments as it is still at an early stage. Recently, Lee et al. [37] found that visualization in VR could help people understand quantities and

* e-mail: cseleni@ust.hk

† e-mail: aoyuwu@seas.harvard.edu

‡ e-mail: wtong@connect.ust.hk

§ e-mail: xxubq@connect.ust.hk

¶ e-mail: zwei302@connect.hkust-gz.edu.cn

|| e-mail: huamin@cs.ust.hk

What animal has the longest neck?

❤️ 1141k 🗨️ 62k

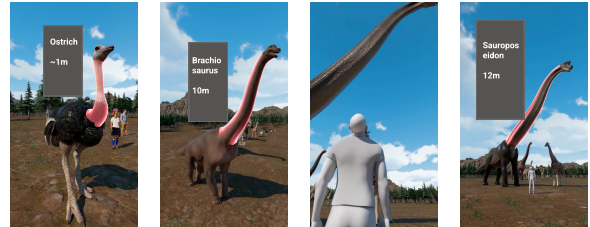


Figure 1: A popular 3D data video on TikTok with rich camera effects. @Permission obtained.

measurements by creating a visceral experience. In this paper, we explore the possibilities of using **VR data videos** for storytelling.

To probe the underexplored use of VR data videos, we start with understanding 3D data videos that display visualizations in 3D space. We are inspired by our observation that many 3D data videos are considerably popular on common channels such as TikTok and YouTube. Those data videos are diverse in designs (e.g., visualization types and styles). They use cinematography techniques in a long take to walk the audience through different data items while manipulating the camera (e.g., camera angles and distances) to enhance the audiences’ experience. For example, Fig. 1 shows a popular 3D data video visualizing the neck length of different animals with rich camera effects such as using low-angle to make the subject look powerful. Such designs echo the form of VR 360 videos [41] that guide the viewers across different scenes such as roller coasters and shark shipwrecks. Thus, we see huge potentials to realize those 3D data videos in VR environments to create a more immersive experience for data storytelling. Towards this vision, we contribute two studies to understand the design factors and design considerations of 3D data videos in screen and VR environments.

In Study 1, we aim to study the design factors of 3D data videos. We analyze 100 most-watched 3D videos on Youtube and Tiktok to derive the design space. Specifically, we draw inspiration from film theory and contextualize them in the context of visualizations through iterative coding. At the top level, our design space describes data videos from two aspects, namely mise-en-scène and cinematography. The former focuses on the stage design (i.e., background and anchors) and the arrangement of actors (i.e., visualizations and their animations). The latter describes cinematographic techniques that are further divided into within-frame (i.e., shot angle and distance) and between-frame (i.e., camera movement and purposes). Centering around the two dimensions, we code categories for each design factor. We report common design patterns and strategies as the results of Study 1.

In Study 2, we seek to understand the design considerations of 3D data videos in desktop and VR environments. We select four design strategies that are assumed to be critical to viewers’ experiences according to our observations and interviews with 13 experts in film, VR, and data storytelling. Then, we conduct controlled experiments to investigate their effects in both desktop and VR environments. 12

participants view those 3D data videos and provide their feedback about their engagement levels, emotional responses, memory, and understanding of the data story. The results show that VR is consistently rated as better than desktop in most cases. This supports our motivation to create a better storytelling experience through VR. We also find that adding an overview significantly improves understandability. Other design strategies, adding anchors, using animations, and varying camera angles, had no significant effect on the stories. Based on the experiment results, we discuss design considerations and implications for creating VR data videos. We then discuss the challenges of designing VR data videos, such as guiding the audience’s attention, reducing motion sickness, and improving aesthetic quality. We make our study material available at <https://immersivedatavideo.github.io/>. To sum up, our contributions are as follows:

- A design space for 3D data videos and an analysis of common design patterns
- A controlled experiment investigating the effects of critical design factors on the viewers’ experiences
- A set of design implications for VR data videos.

2 RELATED WORK

Our study is related to visualization in virtual reality, narrative visualization, and data videos.

2.1 Visualization in Virtual Reality

VR offers extensive space for display and interaction, facilitating analysis with 3D visualizations such as 3D flow maps [64], space-time cube geo-visualization [58], and 3D scatter plots [32, 58] for a wide range of tasks (e.g., comparison and clustering analysis). This field continues to grow with the advancement of visualization techniques in VR. Some research reinvented traditional visualization [33, 40] and interaction techniques [24, 27, 44, 65] to take advantage of the virtual space. Some research re-examined the adaptation of traditional visualization techniques to VR. For example, Whitlock et al. [59] compared the effectiveness of visual encoding channels in terms of graphical perception. Yang et al. [63] explored how to adapt navigation methods in desktop-based visualization to VR environment and compared zooming and overview+detail interactions. Besides conducting evaluation experiments, other work investigated peoples’ interaction patterns with visualization in VR (e.g., [39, 48]). Similar to this stream of research, our work re-examined the effectiveness of 3D visualization designs in data videos in VR and conducted user studies to derive design implications. However, our work focused on a different domain—data videos for data storytelling, in which design strategies are not only related to visualizations but also other storytelling techniques such as camera movement, the setting of scenes, and visual embellishment.

2.2 Narrative Visualization

Narrative visualizations tell data-driven stories engagingly by combining data visualization and storytelling techniques. Initially, Segel and Heer [50] concluded seven genres of narrative visualizations, including magazine style, annotated chart, partitioned poster, flow chart, comic strip, slide show, and film/video/animation. Surrounding these basic genres, follow-up work investigated design factors to enhance the understandability [2, 31, 53], memorability [7, 31], engagement [2], and emotional effects [34, 35], and provided design spaces to increase the expressiveness [36] of data stories. Previous work primarily focused on stories presented on computer screens. Recently, data-driven storytelling in VR started to draw attention from academia and industry [29, 38]. For example, Ren et al. [46] developed a prototype system XRCreator for authoring data storytelling in VR. More work adapted data stories to VR and conducted user studies to derive design implications of data storytelling in VR. Bastiras et al. [5] recreated six stories in VR, which received more

positive feedback than their desktop versions in follow-up user studies. Lee et al. [37] found that VR creates a visceral experience for people to understand the quantities and measurements such as distance, speed, or height. Ivanov et al. [30] re-examined the effects of anthropomorphism and unit visualization techniques and found that they promoted affective responses and personal experiences. Our work complements this field by studying data-driven storytelling in VR 360 video, a popular VR storytelling form [41]. The design factors we investigated come from a systematic analysis of a larger corpus of 100 3D data videos.

2.3 Data Video

Data videos have been a focus of research in data storytelling. Many studies summarized animation techniques. Amini et al. [3] concluded eight animations and their realization in eight chart types. They further found that two design strategies (i.e., the setup animation and pictographs) facilitate the level of understanding and engagement of the participants [2]. Tang et al. [56] summarized animated transition techniques that help stories smoothly switch between scenes. Shi et al. [52] classified animation techniques based on what narrative purposes (e.g., suspense, emphases, ellipsis) they enhanced. Another branch of work focused on creating stories with better structures. Amini et al. [1] labeled sequences of four visual narrative stages (i.e., establisher, initial, peak, and release) in 50 data videos to understand their structures. Similarly, Yang et al. [62] borrowed the traditional narrative structure—Freytag’s Pyramid and explored how to adapt it to data videos. Finally, facing the demand to reduce the burden of creating data videos, Amini et al. [3] developed Dataclips—a template-based authoring tool, and Shi et al. [51] introduced AutoClips that automatically generated data videos from data facts. These works focused on 2D data videos of desktop versions. We studied the design space of 3D data videos and paid more attention to the cinematography techniques such as camera movement in the videos. Importantly, inspired by the similarity of 3D data videos and VR visualizations in the use of 3D scenes and visualizations, we re-examined the design of 3D data videos in VR.

3 STUDY 1: ANALYSIS OF 3D DATA VIDEOS

We aim to understand design considerations for data videos in VR 360 video style by first deriving design factors from 100 highly-viewed 3D data videos online. Next, we introduce our methods for data collection and analysis, followed by describing the result - the design space of 3D data videos.

Study Scope. We focused on 3D data videos that displayed data visualizations in 3D space and walked the audience through each data mark (e.g., in Fig. 1). We excluded data videos with narrations in a film- or documentary-like manner such as *Simulation of a Nuclear Blast in a Major City* [25]. This kind of video has diverse themes but is seldom found in previous research in data videos [10, 52, 61, 62] and our own search. Moreover, data visualizations were often a subsidiary in those videos since non-visualization elements such as narrators, photographs, and archival footage took major parts [8]. We decided to focus on data visualization-centered videos at the current stage. Furthermore, we observed that over 95% of the surveyed data videos visualize two-dimensional datasets. For example, the data visualization in Fig. 1 encodes animal species and neck length. Thus, we discarded videos visualizing more than two-dimensional datasets (e.g., with layered or facet charts) for study manageability. Despite the simple data structure, the videos have variable design dimensions and choices, allowing us to derive many design strategies. Lastly, since we were interested in the VR 360 video style that walks audiences through the scene, we excluded videos without camera movements or location changes, such as the 3D racing bar.

3.1 Data Collection

We collected videos from popular platforms, TikTok and YouTube.

Sampling. We started by searching keywords such as “3D data videos” and “data-driven storytelling”. However, this only generated a few videos from the search result pages, since video creators tended to name their videos by their semantics (e.g., “which animal has the longest neck” in Fig. 1). Nonetheless, the search results helped us identify several KOLs (Key Opinion Leaders) in this field. Their works were recommended on the first few pages by search engines of Youtube and TikTok and they had at least 400K subscribers. Totally, we crawled over 800 videos with 330 Youtube videos and 523 TikTok videos.

Filtering. We ranked the collected videos by the number of likes or views (Youtube only provides the data of views). Then from top to bottom, based on the study scope we mentioned above, we filtered for target videos. We stopped until 100 data videos were reached, which were at a feasible and sufficient scope. Those videos had at least 135K likes or 1.7M views at that time.

3.2 Coding and Analysis Process

We combined top-down and bottom-up approaches to derive the design space. From the top-down perspective, we were inspired by the film theory [6], since it has a formal framework for cinematography techniques that are of virtual importance in 3D data videos but is not studied in 3D visualizations [52]. From the bottom-up aspect, three authors iteratively labeled those data videos. Each video was labeled by at least two authors. We did not label the beginning and ending frames (e.g., showing titles or introducing backgrounds). We held meetings twice a week to discuss and resolve conflicts. The entire labeling process took about two months since it involved identifying camera movements between frames.

Design Space Overview. Fig. 2 provides an overview of the final design space. At the top level, it contains two dimensions mise-en-scène and cinematography. The two concepts are introduced by Bordwell et al. in the book “*Film Art: An Introduction*” [6] (the best-selling and widely accepted textbook of film analysis). They are primary aspects for analyzing film techniques of a shot—the continuous footage between two cuts and the smallest unit of a film. More specifically, **Mise-en-scène** describes the arrangement of four general elements (i.e., stage setting, costumes and makeup, lighting, and staging) that constitute the scene to be shot. **Cinematograph** is the technology of motion-picture photography and covers a wide range of subcategories. Next, we explain our rationale and considerations when deriving the subcategories of the two dimensions.

Mise-en-scène. We mapped the four general elements (i.e., stage setting, costumes and makeup, lighting, and staging) to elements in the 3D data videos. We excluded the lighting as the videos rarely manipulate lights for guiding attention or creating moods as films do. Regarding the stage setting, it is mapped to the 3D background in the data videos. During the bottom-up analysis, we identified two exclusive aspects of the 3D background design. One is the design of the whole scene (whether the background describes a virtual or a real scene), and the other is the use of anchors (independent visual objects to serve as a baseline for audiences to understand the scales of data items). The costumes and makeup and staging decide the appearance and behaviors of actors (the subject of films), which were mapped to the design of data visualization (the subjects of the 3D data videos) and animations happen to it. For visualization design, we coded the mark type and visual encoding channels following Vega-Lite, a popular declarative grammar for specifying 2D visualizations [49]. We coded the mark semantics (i.e., whether data items are represented by abstract graphical marks or concrete real-world objects) and different axis trajectories that decide how 2D visualizations are placed in the 3D environment.

Cinematography. Since little research has explored cinematography in data videos, we started with two fundamental aspects, shot

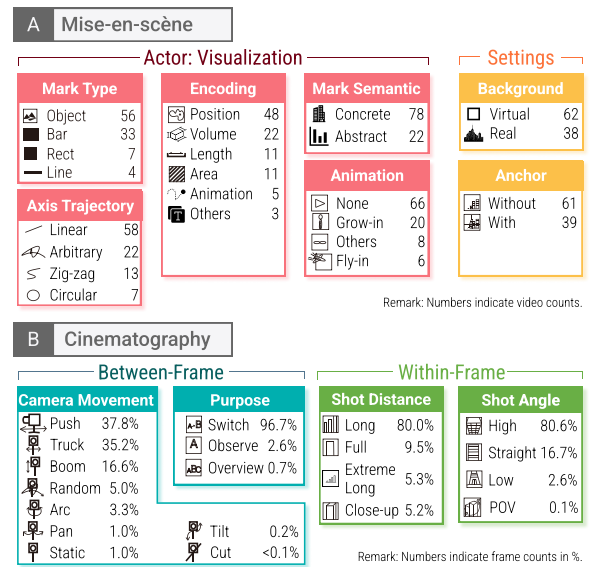


Figure 2: An overview of the design space with statistics.

framing and camera movements, leaving the remaining (e.g., motion speed) to future research. Shot framing and camera movements represent techniques within a frame and between frames respectively. For each frame, shot framing describes the relationship between the camera and subjects in terms of distance and angle. It includes two sub-dimensions: shot angle and shot distance. Camera movements characterize the transition between frames. We further labeled the purpose of camera movements (i.e., whether it is switching between data items, observing a data item, or giving an overview).

3.3 Design Space and Analysis Results

This section explains each dimension of the design space (Fig. 2).

3.3.1 Mise-en-scène

Mise-en-scène describes the visualizations and the background.

Mark Type. Despite that visualization types play an important role in 3D data videos, mark types are limited to three common graphical marks: bars (33%), rectangles (7%), and lines (4%). We are surprised not to find geospatial maps that are common in other visualization genres [1, 53]. Instead, video creators tend to use bars or rectangles for visualizing values of different regions or countries (e.g., Fig. 3 (V5)), emphasizing the rankings instead of geolocation. To sum up, 3D data videos embrace simple visualization designs, which might be for reducing cognitive loads in watching lengthy videos with dense data.

Encoding. Encoding channels are fundamental in data visualizations. Sensibly, position and size channels (including length, area, and volume) are dominant. In 3D settings, the volume channel (22%) is more popular than the length (11%) and area (11%) channel. Notice that a 3D bar encodes the volume channel if all its height, width, and depth vary (e.g., Fig. 3 (V1)). Interestingly, we find the use of animation to encode speed (e.g., Fig. 3 (V6)), which is a new opportunity to bring realistic feelings to immersive 3D environments. Lastly, some cases cannot be well described by encoding channels (3%) since the marks are objects rather than graphical marks. For instance, Fig. 3 (V2) shows the “amount” of different foreign currencies with the same value. To summarize, 3D data videos bring about new opportunities to use volume and animation channels that are less common in 2D settings.

Axis Trajectory. As discussed earlier, we focus on data videos with visualizations encoding two-dimensional datasets. Fundamentally, Those visualizations are in 2D space. Nevertheless, video creators implement several designs to show visualizations in 3D

	V1	V2	V3	V4	V5	V6
Mark Type	Bar	Object	Line	Object	Rect	Object
Encoding	Volume	Others	Position	Length	Area	Animation
Mark Semantic	Abstract	Concrete	Abstract	Concrete	Concrete	Concrete
Axis Trajectory	Zig-zag	Arbitrary	Linear	Zig-zag	Circular	Linear
Animation	None	Without	Grow-in	Without	Grow-in	Fly-in
Background	Virtual	Virtual	Virtual	Real	Virtual	Real
Anchor	Without	Without	Without	With	Without	With

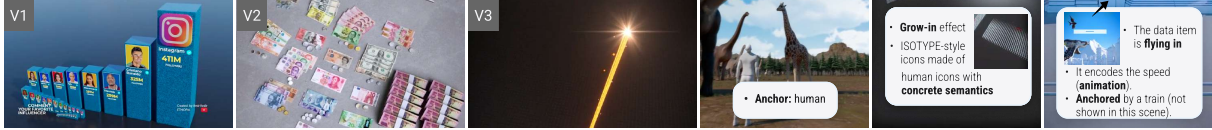


Figure 3: Illustration of the design space under the mise-en-scène dimension with six 3D data videos visualizing (V1) most followed INSTAGRAM accounts [19], (V2) different currencies with a equivalent value [14], (V3) change of world population in history [18], (V4) neck lengths of different animals [28], (V5) sizes of rescue teams from different regions in a disaster [16], and (V6) flying speed of birds [13].

environments. For instance, instead of using linear axes (58%), there are many alternative design choices such as zig-zag (13%, Fig. 3 (V1)), arbitrary (22%, Fig. 3 (V2)), and circular (7%, Fig. 3 (V5)) ones. Those design choices are less common in 2D data visualizations such as timeline visualizations [9]. It suggests considerable efforts from video creators to refresh traditional 2D visualizations in 3D environments to create novel experiences.

Mark Semantics. Compared with 2D settings, 3D environments create a stronger sense of reality. 3D data videos tend not to use abstract graphical elements (22%, e.g., Fig. 3 (V1, 3)), but prefer those with concrete semantic meanings (78%), showing creators’ strong preference for embracing concrete semantics in 3D environments. In addition to real-world objects (V2, 4, 6), another strategy is the use of ISOTYPE visualizations [26]. Fig. 3 (V5) is an example where the rectangle mark is made up of people-shaped pictographs.

Background Semantics. In contrast, video creators less tend to construct backgrounds with concrete semantics, that are, real scenes (38% e.g., natural and city views in Fig. 3 (V4, 6)). On the contrary, they often use virtual backgrounds (62%, e.g., Fig. 3 (V1, 2-3, 5)). It can be due to the difficulties and efforts needed to build real, semantic backgrounds, or that creators attach greater importance to visualizations than backgrounds.

Anchor. We note the use of anchors (39%) that are similar to props in films. Specifically, those anchors provide baselines for audiences to compare data scales to gain a more realistic and intuitive understanding of data. For instance, Fig. 3 (V4) puts a human model to help audiences understand the length of animal necks, and Fig. 3 (V6) shows a running train to immerse audiences in feeling the speed of birds. Such anchors are particularly useful when the visualizations encode real-world measurements such as speed and height.

Animation. Animation design has been a research focus in data visualizations [23, 52, 53, 56]. Nevertheless, animated visualizations are uncommon in 3D data videos, i.e., 34% of videos animate the visualizations. Used animation techniques are limited to basic ones such as growing-in effects (20%, Fig. 3 (V3)) and flying-in effects (6%, Fig. 3 (V6)). In contrast, 3D data videos animate the camera using cinematography techniques, i.e., only 1.0% of video frames have static camera movements. We consider cinematography a distinct aspect of 3D data videos and detail it in the next section.

3.3.2 Cinematography

Cinematography concerns each frame and transitions between them. To label it, we sampled one frame for every second in the video. We labeled the start and end frames with the same camera movement and purpose. Note that different camera movements could be used together, such as tilting up and booming up at the same time. We labeled the shot distance and angle for each frame.

Purpose. We start with analyzing the purpose dimension, as it represents the types of content a camera is shooting for. Most

frames in our corpus present the switching (96.7%, e.g., Fig. 4 (V3)) between data items. We observe that videos reveal data items in a continuous movement most time, seldom staying around a visual mark to observe (only 2.6% of frames were for observing a single visual mark, e.g., Fig. 4 (V2)). As 89% of videos interweave between switching and observing data items, this suggests that most videos apply the strategy of only carefully observing a small number of items (e.g., the biggest ones and outliers) for emphasis. Overview (e.g., Fig. 4 (V1)) takes the smallest percentage of frames (0.7%), as only 29 videos used it, and 26 of them used it in the last frame.

Camera Movement. We reference the classification of camera movement in film [55]. Specifically, the camera movements push, truck, boom, and arc move the camera on a predefined truck without rotating it. Pushing moves the camera close to or further away from an object. Trucking moves the camera left or right and booming moves the camera up or down. Arcing moves the camera around a subject in an arc pattern. The camera movements pan and tilt (e.g., Fig. 4 (V2)) direct a camera horizontally left or right and vertically up or down, respectively, without changing its position. Finally, static, random, and cutting are special cases. The name static speaks for its meaning. Random corresponds to irregular movements such as a slight shaking of the camera. Cut means a sudden change from one scene to another without moving the camera in continuous time and space. Next, we analyze camera movements and their purposes together to understand video creators’ design choices.

For the purpose **switch**, push (38.32%), truck (35.22%), and boom (16.04%) are widely used. When too large (small) or high (short) items show, the camera pushes out or booms up to reveal them clearly. Otherwise, a simple truck is good enough. In fewer situations, videos use arc (4.19%), pan (3.37%), random (1.68%), and tilt (1.01%). A close investigation reveals that pan and arc are used together with push or truck to shoot items at turning points in trajectories. Sometimes, the movements arc, pan, random, and tilt are for adding variations or changing perspectives. The use of tilt interests us the most. When switching to significantly high or large data items, some videos use tilt to enhance the audience’s feeling of height, creating a sense of owing. When **observing** certain items, most videos use static (31.53%), random (19.88%), push (16.60%), and arc (11.50%). While stopping the camera for observation is an intuitive treatment, more videos prefer to use random, push, and arc to mimic human eyes observing an item from different directions, adding energy to the shot [55]. As for the purpose **overview**, the videos use push (43.92%), boom (15.35%), and static (15.14%) most. Naturally, to include all data items in the shot, the camera pushes out and booms up. Sometimes, videos need to use arc (13.65%) and pan (9.38%) to face the camera to the front of the data items. Finally, many videos freeze for about 1-3 seconds in the overview.

Shot Distance. Shot distance is originally defined by how much of the environment and characters (e.g., body parts from the knees

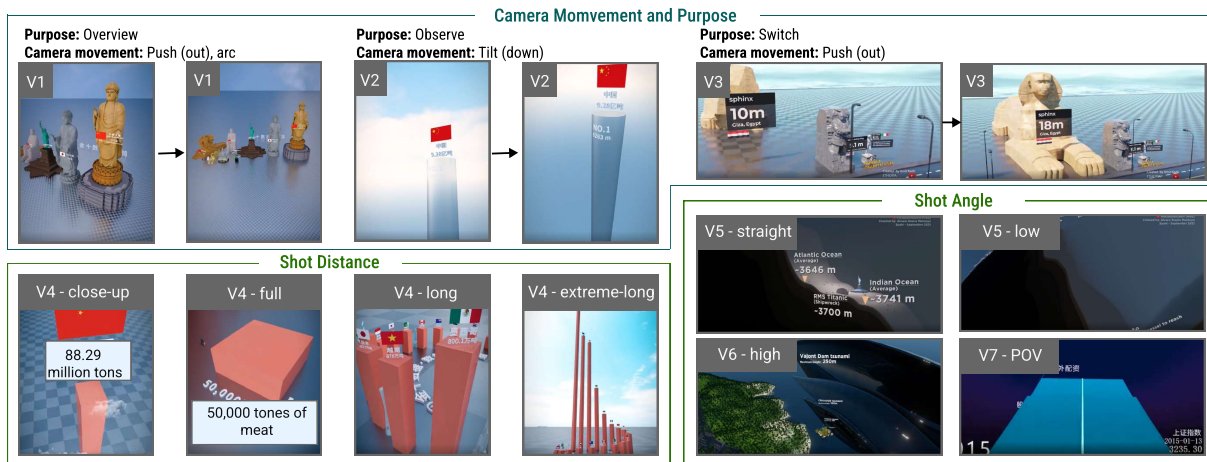


Figure 4: Illustration of the design space under the cinematography dimension with seven 3D data videos visualizing (V1) heights of World famous sculptures [12], (V2) productions of steel of top 10 countries [15], (V3) sizes of tallest statues [20], (V4) the consumption of meat in various countries [11], (V5) depth of ocean [43], (V6) biggest tsunamis [42], and (V7) trend of the stock index [21].

up) are in the shot [54]. We adapted the definition based on how many visual marks are in a shot and classified them into close-up (parts of a visual mark), full (one visual mark), long (two or more visual marks), and extreme long (all visual marks) shots (e.g., Fig. 4 (V4)). The long shot (80%) is mostly used to help compare visual items next to each other. Full (9.5%) and close-up (5.2%) shots are used less, and extreme long shot (5.3%) naturally comes with showing the overview of the whole visualization. Most videos (47) have both cases of increasing and decreasing the camera distance for observing some items closely. Some videos (37) increase camera distance (from full/close-up to long to extreme-long) throughout the whole video. Only 16 videos never changed shot distance.

Shot Angle. We concluded four types of camera angles: high (80.6%), straight (16.7%), low (2.6%), and POV (0.1%). The camera angles, high, straight, and low, follow their literary meanings. POV represents a special case of seeing a chart in a first-person view. We found that most videos use only one angle type, high angles (39), straight angles (14), or low angles (1) the whole time. For other videos (54), they use one angle as the dominant one (straight or high) and change the angle to straight, low, or high for observing large items or all items in the later part of the videos.

4 STUDY2: EXPERIMENTS WITH VR VIDEOS

Study 2 explored the possibility of adapting design strategies from 3D data videos to VR data videos. We evaluated four design strategies in VR that were selected based on our observation from Study 1 and consultants with 13 experts in VR, data storytelling, and film.

4.1 Pilot Study: Expert Interview

To make our study manageable, we hoped to select a subset of design factors that were potentially more important in user experiences. We conducted an interview study with experts in VR, film, and data storytelling. Specifically, we recruited 13 experts in VR (E1 - E5, with 1 to 4 years of experience), film (E6 - E9, with 2 to 25 years of experience), and data storytelling (E11 - E13, with 2 to 6 years of experience). We first introduced our design space to the experts and gave them full access to the video gallery on our website. Then, they rank the design strategies from an initial list of comparisons from four aspects: understandability, memorability, engagement, and emotional effect. As mentioned in Section 2.2, these metrics were commonly used for evaluating data stories in previous research. In the initial list, each comparison had at least two design strategies from one or a combination of dimensions in our design space. These strategies were popular in the corpus, or we found to bring different viewers' experiences. When they found the effects of a strategy

depended, they gave their explanations. For example, E10 marked the effects of "using anchors" as it depends and explained that when people already could feel the quantity of data, anchors could become distractions. Afterward, we asked experts what factors they found most important for viewers' experiences. Moreover, we asked VR experts whether these strategies could be adapted to VR and what should be paid attention to for the adaptation. They provided instructive suggestions for our implementations of VR videos in the following experiments.

Based on the interview results, we filtered out comparisons with strategies that were equally valued. For example, six participants thought that different trajectories would make limited differences, and three participants suggested that these trajectories facilitate comparing items at different directions and levels (e.g., comparing two adjacent items in a linear trajectory, comparing two lines of items in a zigzag trajectory). We also eliminated dimensions that were not mentioned by participants as important, as well as design strategies that had been explored in the work by Lee et al. [37] such as the realism of the scene and data visualizations. Finally, we selected four comparisons: (1) adding **anchors** or not, (2) keeping camera angles the same versus **changing camera angles**, (3) using **animation** or not, and (4) showing an **overview** at the end or not.

4.2 Study Design and Procedure

We conducted a controlled experiment to evaluate the above four design strategies.

4.2.1 Experiment Material

We implemented the cases with two datasets and the two most basic chart types (bar charts and line charts) in our corpus. We created a control case for each chart and another four experimental cases using the strategies (Fig. 5).

The Bar Chart Case. The bar chart [17] showed the top 15 countries with the most tea production. In the *control case*, the camera was kept at a high angle to show each bar. It moved from left to right, boomed up when a bar was too high to reveal its top part, and stopped at the last 3D bar. Next, we introduce how we modified the control case to use the four design strategies. For the case having *anchors*, we added 3D models of humans, cars, and famous buildings (e.g., the Eiffel Tower). For the case that changed *camera angles*, we followed the design of the original video. The camera shot each bar at a low angle at the beginning, changed to a high angle in the middle, and tilted up and down when revealing the two highest bars. This was also common in other videos. In the case that used *animation*, each bar grew up when the camera moved

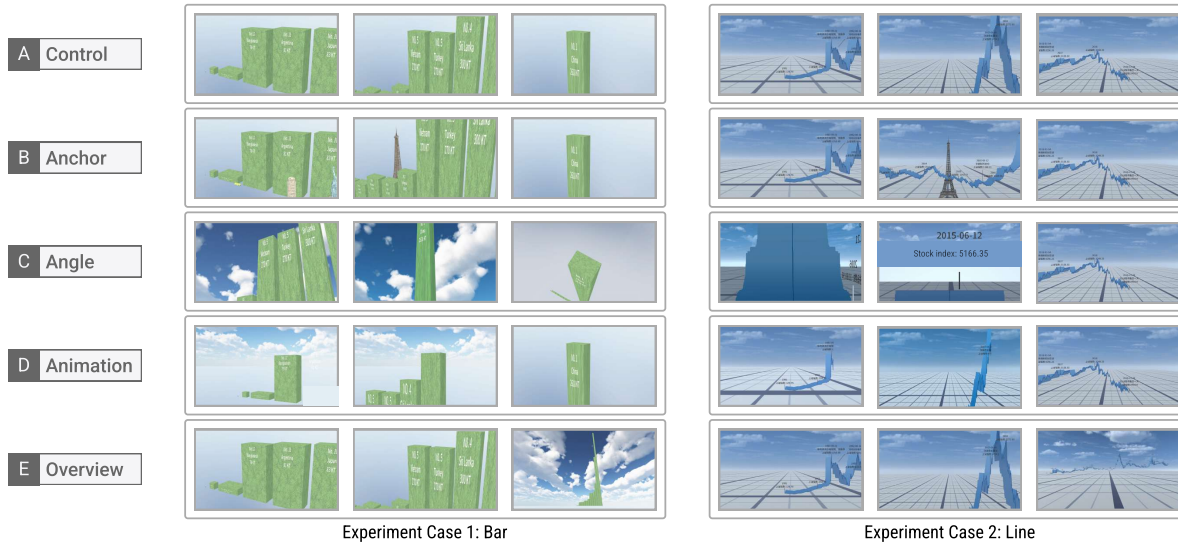


Figure 5: This figure shows the keyframes of the videos in five conditions: (a) the control cases, (b) adding anchors, (c) varying camera angles, (d) using animation, and (e) showing an overview at the end.

close to it. Finally, for the case that showed an *overview*, we added an extreme long shot at the end. Besides the above modifications, the other settings of the four cases (e.g., camera speed, the scene, lighting) were the same as the control case.

The Line Chart Case. We used the data of the Shanghai Composite Index from 1991 to 2018 [21]. The *control case* had the same design as the control case of the bar chart. A camera moved from left to right to gradually show a 3D line chart without any anchor or animation. We used similar *anchors* in the bar chart, except that we mapped 10 indexes to one-meter height. As for changing *camera angles*, we changed the point of view from “third-person” to “first-person” in the two biggest peaks that happened from 2006 to 2008 and from 2015 to 2016. The audience would have a similar view that happened when riding a roller coaster. At last, for the case with an *overview*, we used the same treatment with the bar chart.

Adapting Videos in VR. As suggested by experts in VR, we carefully designed the camera movements in VR. Following the form of VR 360 videos [41], while we have predefined camera movements, the participants could freely move their heads to look around. The camera movement speed was set at not too slow, which might cause boredom, nor too fast, which might not show data clearly. We asked three external researchers to experience our cases and two of them reported motion sickness. Following previous research on motion sickness in VR [45], we took the following strategies. First, when the camera needed to stop and change direction, we use the ease-in and ease-out effects. Second, when the camera changed large degrees (e.g., switch between the first-person and the second-person view), the movement speed was set fast to mimic the effect of jumping between scenes, which caused less sickness than a smooth transition but still could help the audience orientate themselves. Finally, when showing an overview, instead of pushing back the camera continuously, the scene was changed instantly.

4.2.2 Mixed Design

We created 10 VR and 10 desktop videos. We hope to understand the effects of the four design strategies and their differences in VR and desktop. However, we faced difficulties in recruiting enough participants to conduct in-lab studies during the pandemic and it would be a huge burden for each participant to see all cases if we took a within-subject experiment design. Therefore, to reduce the burden of participants and still make a reasonable sample size, we took a mixed design that combined the between-subject and

within-subject design [4]. The between-subject factor was the chart type (i.e., bar chart and line chart), and the within-subject factors were the device (i.e., VR and desktop) and the case (i.e., control case and experimental case). For each comparison between an experimental case and a control case, half of the participants would see bar charts, and another half of the participants would see line charts. As the experimental cases of each chart shared the same control case, participants only needed to watch the control cases of each chart one time. Participants would see both control and experimental cases in VR and desktop. To reduce the learning effects, one participant would see line charts in two comparisons and see bar charts in another two comparisons. Under this mixed design, we had a sample size of 12 participants for each comparison between the control and the experimental cases. Every participant watched 12 videos in total.

4.2.3 Experiment Procedure

We recruited 12 participants (six females) from a university. Most of them were postgraduate students. Only four of them were familiar with VR, and the others had few or no VR experiences. In the beginning, we introduced the concept of VR data videos. They were told that the video was designed as they were sitting on an aircraft to have sightseeing. The initial direction they faced was always the direction the aircraft faced. Every participant watched both the desktop and VR versions of six cases in random order. Among the six cases, two were control cases of the line chart and the bar chart. Two were experimental cases with the line chart from two comparisons. The remaining two were experimental cases with the bar chart from the other two comparisons. We made sure to have all possible combinations of the chart type and the types of the two comparisons. We conducted the experiments in a large open space, using an Oculus quest2 connected to a Macbook Pro with an Apple M1 chip and 8 GB of RAM.

When finishing watching a video, every participant rated the understandability, memorability, engagement, and emotional effect of the video on a 7-point Likert scale and gave comments. We followed methods in [36, 47] to measure the four metrics through subjective ratings, which was the most feasible for our task. While all other metrics have been evaluated subjectively in previous research about visualization evaluation, the evaluation of memorability often involves asking participants to recall the visualization [47]. It was not suitable for our case as participants saw the same story several

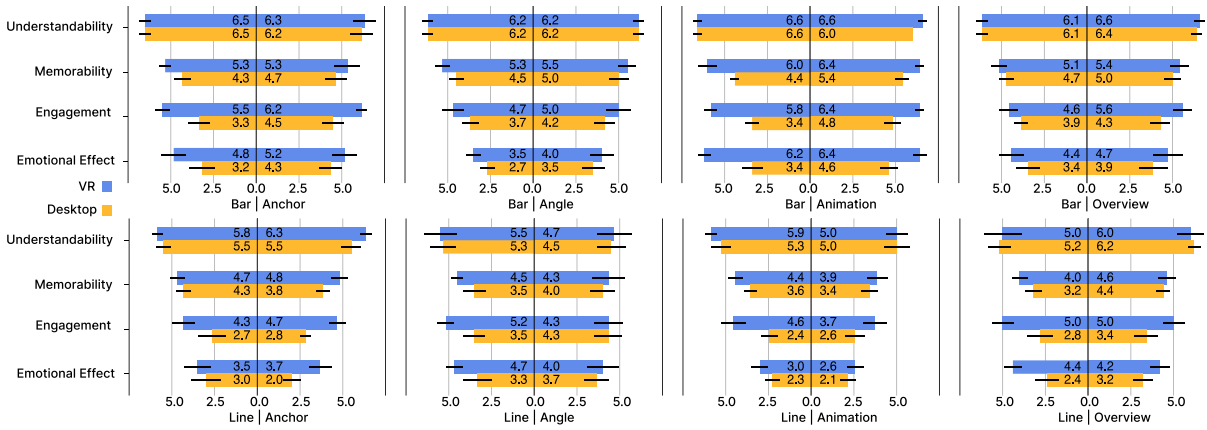


Figure 6: The chart shows the average scores with standard errors for each comparison. The upper and lower sections show data of the bar chart cases and the line chart cases respectively. In each grouped bar chart, the left part shows the data of the control case, and the right part shows the data of the experimental case. Blue bars represent data of the VR cases, and yellow bars represent data of the desktop cases.

times. Thus, measuring understandability and memorability through subjective ratings instead of objective tasks (e.g., visualization tasks for testing comprehension) was a compromise to avoid the learning effects. Nonetheless, we emphasized to the participants that “memorability” was not about how many numbers they could remember exactly from the video but whether the video could leave a long-last impression in their memories. Thus, participants’ responses were still valuable references for achieving the goal of these videos in promoting science popularization. Lastly, we interviewed participants to gather feedback on the following questions: (1) Did and how did VR bring them different experiences? (2) What are the design strategies they appreciated (if any)? (3) What are the potential improvements in these cases?

5 USER STUDY RESULT

We analyzed participants’ ratings on each metric for each comparison between the control and the experimental case independently. We used a Mixed ANOVA in SPSS Statistics with chart type being the between-subject factor, the type of device and case (i.e., control or experimental case) being within-subject factors, and the score on each metric being dependent variables. Next, we report on both statistical analysis results (see Fig. 6) and qualitative feedback from the participants for each comparison. In the results, η^2_{partial} means effect size—the magnitude of the difference between compared groups, and MD represents the mean difference between two groups based on estimated marginal means.

Anchor. Participants gave significantly higher scores of emotional effect ($P < 0.05, \eta^2_{\text{partial}} = 0.489, MD = 1.167$), engagement ($p < 0.001, \eta^2_{\text{partial}} = 0.715, MD = 1.833$), and memorability ($p < 0.001, \eta^2_{\text{partial}} = 0.717, MD = 0.750$) when using VR than desktop, but not the case for the understandability ($p > 0.1$). The engagement of videos with anchors had a weak advantage over those without ($p < 0.1, \eta^2_{\text{partial}} = 0.310, MD = 0.583$). However, having anchors did not significantly affect the understandability, memorability, or emotional effect. There was a weak crossover interaction between chart type and the use of anchors ($p < 0.1, \eta^2_{\text{partial}} = 0.288$). When participants viewed the bar chart, the average score of the emotional effect of cases with anchors was a little higher than those without ($p < 0.1, MD = 0.75$). On the contrary, for line charts, having anchors drove a smaller mean score of the emotional effect ($p > 0.1, MD = -0.417$). Participants reported that anchors did not facilitate understanding mainly for two reasons. First, for the line chart, four participants ($P5-7, 11$) found that they could not relate the anchors to the abstract stock index data. Second, two participants ($P4, 7$) had seen the buildings in real life, and the buildings were like miniaturized models in VR. The unrealistic experience brought

no benefit to understanding.

Angle. We found no significant positive effects of changing angles. VR was significantly better in terms of memorability ($p < 0.05, \eta^2_{\text{partial}} = 0.368, MD = 0.667$) but no significant effect in understandability ($p > 0.1$). Surprisingly, we found a weak interaction effect of device type and the use of angles for metrics emotional effect ($p < 0.1, \eta^2_{\text{partial}} = 0.286$) and engagement ($p < 0.1, \eta^2_{\text{partial}} = 0.294$). When there was no change of angle, VR was better than desktop in terms of emotional effect ($p < 0.05, MD = 1.083$) and engagement ($p < 0.05, MD = 1.333$). On the other hand, with angle changes, VR was not significantly better. We found that the weakened effectiveness of VR with camera changes might be a result of two factors. First, four participants ($P2, 10-12$) reported a slight sickness. Second, we observed situations where when the camera tilted up in the bar chart cases, two participants were looking aside and missed the moment. We received diverging comments from participants on changing angles. Three participants ($P1, 9, 11$) indicated that the line chart in a “first-person” view might cause confusion with a loss of the big picture. However, seven participants ($P3, 5, 7, 9-11$) agreed on the positive effects of changing angles in engagement by bringing variations and enhancing the sense of depth and height.

Animation. There was no main effect of the use of animation. We still observed the positive effect of using VR in memorability ($p < 0.05, \eta^2_{\text{partial}} = 0.614, MD = 0.971$), engagement ($p < 0.001, \eta^2_{\text{partial}} = 0.878, MD = 1.821$), and emotional effect ($p < 0.001, \eta^2_{\text{partial}} = 0.889, MD = 1.436$), but not for understandability ($p > 0.1$). Participants had two opposite attitudes toward the use of animation. Four participants ($P6, 8-9, 11$) disliked animation and thought processing both the animation and camera movements brought burdens to them. Moreover, in cases without animation, VR provided a wide view so that they could preview the whole chart by turning around. Otherwise, they had to wait for the animation to finish, which hindered their understanding. On the contrary, six participants ($P3, 5, 9-10, 12$) appreciated that animation guided their attention and commented that the grow-up animation of data items worked like a double-encoding, increasing their sense of the data.

Overview. Compared with desktop, VR was significantly better in engagement ($p < 0.05, \eta^2_{\text{partial}} = 0.531, MD = 1.450$) and emotional effect ($p < 0.05, \eta^2_{\text{partial}} = 0.531, MD = 1.214$), and had a weak advantage in memorability ($p < 0.1, \eta^2_{\text{partial}} = 0.306, MD = 0.464$), but not for understandability ($p > 0.1$). We found significantly positive effects of overview in understandability ($p < 0.05, \eta^2_{\text{partial}} = 0.458, md = 0.679$), and weak positive effects in memorability ($p < 0.1, \eta^2_{\text{partial}} = 0.322, md = 0.593$), engage-

ment ($p < 0.1, \eta^2_{\text{partial}} = 0.280, md = 0.507$), and emotional effect ($p < 0.1, \eta^2_{\text{partial}} = 0.285, md = 0.329$). Participants agreed on the effectiveness of an overview that provided a comprehensive view of the data and evoked affecting responses by enhancing the perception of differences between data items.

6 DISCUSSION

Next, we discuss the implications and the limitation and future work.

6.1 Design Factors and Implications

Give references that are semantically related to data. Many videos in our corpus had concrete backgrounds, visual marks with realistic textures and shapes, and real building models as anchors. In expert interviews, 11 experts emphasized the importance of adding realism to VR videos for immersion. A good example is the video in which water from the sea rose like mountain peaks, forming visual marks to compare the heights of tsunamis. However, experts (*E2-3, 8-9, 11-13*) proposed that the background, visual marks, and anchors could distract the audience or even harm their understanding of the story if they were not semantically related to data. For example, the case that compared the neck lengths of different creatures by presenting their models in a mountain view received negative comments. Experts found the mountain view was irrelevant, causing confusion. In that case, an abstract background was preferred. Similarly, the user study participants (*P2, 8, 12*) reported the difficulties of connecting building models to data in the line chart. Furthermore, be cautious of visual marks or anchors that are partially realistic due to scale transformation or missing essential details. For instance, due to the scale transformation, two participants (*P2, 9*) were confused by the differences between the building models as anchors in reality and VR. When watching a video comparing the numbers of bones of creatures, one expert (*E10*) found that the bones confusedly had the same shapes.

Camera angle: Create a comfortable and guided trip. 10 experts remarked highly on the variance of camera angles for providing new perspectives and evoking emotions. However, applying it requires much effort into reducing motion sickness and clearly conveying the narrative intention. The movements and rotations of the camera in the line chart case caused slight motion sickness for some participants (*P2, 8, 10-12*). While we applied jump cuts for moments with a great degree of changes in the camera angle, the rest small changes still caused motion sickness. Data storytellers could consider other methods such as restricting the field of view [22]. Moreover, not all participants appreciated the change of angles when they could not understand its narrative intention. For example, the first-person point of view in the line chart was to create a feeling of going up and down with the trend of the stock market to show its uncertainty and risks. However, without narrations to indicate the message, a few participants could not fully understand the video. Sometimes, participants could miss the intention of the narrative when their own sights conflicted with the direction of the camera. While most participants in the experiment went back to the initial direction to follow the story, two participants (*P6, 11*) reported that they felt lost at some moments. This could be avoided by deliberately guiding the audience's attention with visual cues.

Animation: Balance the reader-driven and author-driven storytelling and reduce the cognitive load. The participants' disagreements on the effectiveness of animation came from their preferences for reader-driven and author-driven storytelling [50]. While some participants (*P3-5, 9-12*) valued the animation for guiding their attention and enhancing the sense of data, others indicated that it restricted their ability to preview the data in VR. To balance the reader-driven and author-driven storytelling, we suggested the following alternative designs that we observed from our corpus. First, a moving visual cue, such as a dot sliding on a line chart, could be applied to guide attention. Second, video creators could add

contours of visual marks while keeping a grow-in animation. At last, other animated effects such as blinking, shaking, and bouncing are also worth trying. When selecting animation techniques, an important consideration is how to reduce the cognitive load of the audience. Five experts (*E2, 10-13*) suggested that animation should have semantic relation with the data itself to avoid distractions. For example, for the video comparing the sizes of craters, an animation of a rock falling down from the sky was more appropriate than a fade-in animation. Second, adding animations could increase the scene's visual complexity too much, which also concerns other visual elements. For instance, one expert (*E10*) commented that the video comparing the numbers of bones of different creatures already had a complex background and visual marks, and the animation of each bone flying made the scene too complex.

Camera movements and trajectories: Consider data patterns and perspectives. Moving the camera along a trajectory with variations of the camera's angle and distance to progressively reveal data was the essential method to manipulate the view experience. All experts mentioned that choosing camera movements and trajectories should depend on data patterns. The first matter is whether there are significant differences between data. Eight experts (*E1, 4-7, 10-12*) suggested that movements such as boom and tilt were primarily for revealing differences between data items. Similarly, five experts (*E2, 10-13*) indicated that the linear trajectory could better emphasize the differences between data items than circular and zig-zag trajectories. Secondly, storytellers could choose strategies based on the set of data items they focus on. Four experts (*E2, 10-11, 13*) mentioned that the zig-zag and circular trajectories were more suitable for comparing multiple items, while the linear trajectory focused more on the nearest item. Pushing and arcing movements can also be used to switch between showing details of one data item and comparison between multiple ones. Finally, choosing movements and trajectories could also come from the purpose of entertainment. Five experts (*E6-10*) suggested arbitrary trajectories and camera movements could prevent making the audience bored.

6.2 From Desktops to VR

All participants in the experiments praised VR for its wide view and immersiveness. Nine participants (*P1-4, 6-8, 10, 12*) mentioned that being able to look forth and back in VR helped them review and preview data which facilitated understanding and memory. Eight participants (*P3, 5-8, 9-11*) reported that "moving up and down" along the data trend and viewing the data items like high buildings or mountains in real life created a sense of "wow" in VR. The statistical results also showed that participants rated VR as being significantly better than desktop in terms of memorability, engagement, and emotional effects, without hindering the understandability. Despite its advantages, the following attributes of VR could make it difficult to achieve the same effects of some design strategies in VR as desktop-based videos do. First was the issue of motion sickness brought by camera movements, as mentioned in Section 6.1. Second, the function of freely changing the view direction in VR made it hard to let the audience precisely follow and catch the intention of camera movements. Moreover, some participants (*P1, 4, 7, 10-12*) reported that they were not very used to the lower resolution in VR, having negative impacts on the aesthetic quality. Finally, three participants (*P2, 4, 7*) felt that the speed of movement was slower in VR than on desktops because the wider view of VR made less obvious visual changes in the scene when the camera moved.

6.3 Limitation and Future Work

Other design dimensions. Our design space does not concern the design of sound, music, and speed of camera motion that are important to create the mood and atmosphere [61]. Furthermore, we excluded film- or documentary-like 3D data videos that have complex narratives and visual elements to create an engaging storytelling

experience. They are also very worth to be explored in the context of VR. Finally, due to the insufficiency of VR data stories corpus, we drew inspiration from 3D data videos displayed through 2D screens. Therefore, some unique characteristics of VR were not considered such as the interaction and the position of elements in a larger 3D space relative to the audience. As concluding related strategies require other frameworks, we leave it for future work.

More visualization types. The videos in our corpus used the most basic data visualizations. Visualization types such as network, matrix, flow, and maps were under exploration. Moreover, we focused on videos with two-dimensional datasets. One film expert (E5) mentioned that with one more dimension to code information, the purposes and effects of camera movements could be quite different from that in the current videos. For example, for a data chart that encodes data in all three dimensions, changing the orientations and directions of the camera reveals new information and perspectives. Camera movements could achieve more dramatic effects such as suspense and twist. Future work is required to collect or create more complex cases.

Expanding the scope of the experiments. Due to limited resources, we applied a mixed design and let participants see the same datasets several times to increase the sample size. While we randomized their watching orders to minimize the learning effects, we could not entirely eliminate its potential influences on the quality of participants' feedback in the trials in the later part of the experiments. Moreover, we focused on a subset of design strategies without considering the combined effects of different design dimensions. In the future, we plan to increase the scope of our experiments in terms of both increasing the number of subjects and design strategies.

Tools for creating 3D VR video. We created the VR video with Unity, the mainstream software for developing VR applications. We found that the whole process was far more time-consuming than creating 2D visualizations with animations. For example, setting the path and turning points for camera movements and rotations in a 3D space required significant attempts and adjustments. Coordinating animation and camera movements could be even harder. Creating 3D visual marks with semantic texture and shape, like creating 2D pictographs, is also challenging. Future tools could provide (semi-)automatic supports [60] for facilitating the setting of the camera and animation based on data patterns and templates for creating expressive 3D data visualizations.

7 CONCLUSION

This study explored the possibilities of adapting 3D data videos designs to VR 360 data videos. We proposed a design space by analyzing 100 highly-watched 3D data videos. We then evaluated four design strategies (i.e., anchors, animation, camera angles, overview) from four aspects (i.e., understandability, memorability, engagement, and emotional effect) in both VR and desktop data videos. The result showed that VR was rated as more memorable, engaging, and evoking more emotional effects than desktops. For design choices, giving an overview significantly improved the understandability of data videos, while the other three design strategies did not significantly affect the quality of videos. While most participants appreciated them, their effectiveness depends on more factors such as their semantic relations to data and participants' preferences for author-driven and reader-driven storytelling experiences. We provided a set of design implications for better applying these design strategies. Future work could expand the experiment scope, explore VR data videos with more complex designs, and develop authoring tools.

ACKNOWLEDGMENTS

The authors wish to thank all the reviewers, experts, and experiment participants for their valuable feedback.

REFERENCES

- [1] F. Amini, N. Henry Riche, B. Lee, C. Hurter, and P. Irani. Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1459–1468, 2015.
- [2] F. Amini, N. H. Riche, B. Lee, J. Leboe-McGowan, and P. Irani. Hooked on data videos: assessing the effect of animation and pictographs on viewer engagement. In *Proceedings of the International Conference on Advanced Visual Interfaces*, pp. 1–9, 2018.
- [3] F. Amini, N. H. Riche, B. Lee, A. Monroy-Hernandez, and P. Irani. Authoring data-driven videos with dataclips. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):501–510, 2016.
- [4] A. P. Association. Apa dictionary of psychology - mixed design. <https://dictionary.apa.org/mixed-design>, 2018.
- [5] J. Bastiras and B. H. Thomas. Combining virtual reality and narrative visualisation to persuade. In *International Symposium on Big Data Visual Analytics*, pp. 1–8. IEEE, 2017.
- [6] D. Bordwell, K. Thompson, and J. Smith. *Film art: An introduction*, vol. 9. McGraw-Hill New York, 2010.
- [7] M. A. Borkin, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.
- [8] J. D. Bradbury and R. E. Guadagno. Documentary narrative visualization: Features and modes of documentary film in narrative visualization. *Information Visualization*, 19(4):339–352, 2020.
- [9] M. Brehmer, B. Lee, B. Bach, N. H. Riche, and T. Munzner. Timelines revisited: A design space and considerations for expressive storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2151–2164, 2016.
- [10] M. Conlen, J. Heer, H. Mushkin, and S. Davidoff. Cinematic techniques in narrative visualization. *arXiv preprint arXiv:2301.03109*, 2023.
- [11] J. Data. Compare the consumption of meat in various countries. <https://www.douyin.com/video/6651416280689593612>, 2019.
- [12] J. Data. Compare the heights of world famous sculptures. <https://www.douyin.com/video/6670696076120444167>, 2019.
- [13] J. Data. The fastest birds in the world. <https://www.douyin.com/video/6703807324387052803>, 2019.
- [14] J. Data. How much foreign currency can be exchanged for 100 rmb. <https://www.douyin.com/video/6763560468951878920>, 2019.
- [15] J. Data. Top 10 countries producing most steel in 2018. <https://www.douyin.com/video/6668247345450323203>, 2019.
- [16] J. Data. The number of medical staff who went to hubei to assist across the country. <https://www.douyin.com/video/6805855459669593357>, 2020.
- [17] J. Data. Top 20 countries with the most tea production. <https://www.douyin.com/video/6870324535762111757>, 2020.
- [18] J. Data. History of changes in world population. <https://www.douyin.com/video/6985462271614995719>, 2021.
- [19] R. Data. Most followed instagram accounts comparison in 3d. <https://www.youtube.com/watch?v=irBTlNXJrfo>, 2021.
- [20] R. Data. Tallest statue size comparison. <https://www.youtube.com/watch?v=YEPJdkZRdJU>, 2021.
- [21] DATA-VIZ.cn. Stock market crash in 28 years. <https://v.qq.com/x/page/10818ao7dtb.html>, 2018.
- [22] A. S. Fernandes and S. K. Feiner. Combating vr sickness through subtle dynamic field-of-view modification. In *IEEE Symposium on 3D User Interfaces*, pp. 201–210. IEEE, 2016.
- [23] T. Ge, Y. Zhao, B. Lee, D. Ren, B. Chen, and Y. Wang. Canis: A high-level language for data-driven chart animations. In *Computer Graphics Forum*, vol. 39, pp. 607–617. Wiley Online Library, 2020.
- [24] T. Gesslein, V. Biener, P. Gagel, D. Schneider, P. O. Kristensson, E. Ofek, M. Pahud, and J. Grubert. Pen-based interaction with spreadsheets in mobile virtual reality. In *IEEE International Symposium on Mixed and Augmented Reality*, pp. 361–373. IEEE, 2020.
- [25] N. Halloran. Simulation of a nuclear blast in a major city. <https://www.youtube.com/watch?v=Z3RzNEzJyzo&t=8s>, 2020.
- [26] S. Haroz, R. Kosara, and S. L. Franconeri. Isotope visualization:

- Working memory, performance, and engagement with pictographs. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1191–1200, 2015.
- [27] Y.-J. Huang, T. Fujiwara, Y.-X. Lin, W.-C. Lin, and K.-L. Ma. A gesture system for graph visualization in virtual reality environments. In *IEEE Pacific Visualization Symposium*, pp. 41–45. IEEE, 2017.
- [28] W. is the animal With the longest neck. Jinshi data. <https://www.douyin.com/video/7015099278217497892>, 2021.
- [29] P. Isenberg, B. Lee, H. Qu, and M. Cordeil. Immersive visual data stories. In *Immersive Analytics*, pp. 165–184. Springer, 2018.
- [30] A. Ivanov, K. Danyluk, C. Jacob, and W. Willett. A walk among the data. *IEEE Computer Graphics and Applications*, 39(3):19–28, 2019.
- [31] H.-K. Kong, W. Zhu, Z. Liu, and K. Karahalios. Understanding visual cues in visualizations accompanied by audio narrations. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.
- [32] M. Kraus, N. Weiler, D. Oelke, J. Kehrer, D. A. Keim, and J. Fuchs. The impact of immersion on cluster identification tasks. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):525–535, 2019.
- [33] O.-H. Kwon, C. Muelder, K. Lee, and K.-L. Ma. A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(7):1802–1815, 2016.
- [34] X. Lan, Y. Shi, Y. Wu, X. Jiao, and N. Cao. Kineticcharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):933–943, 2021.
- [35] X. Lan, Y. Shi, Y. Zhang, and N. Cao. Smile or scowl? looking at infographic design through the affective lens. *IEEE Transactions on Visualization and Computer Graphics*, 27(6):2796–2807, 2021.
- [36] X. Lan, X. Xu, and N. Cao. Understanding narrative linearity for telling expressive time-oriented stories. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.
- [37] B. Lee, D. Brown, B. Lee, C. Hurter, S. Drucker, and T. Dwyer. Data visceralization: Enabling deeper understanding of data using virtual reality. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1095–1105, 2020.
- [38] B. Lee, T. Dwyer, D. Baur, and X. G. Veira. Watches to augmented reality: Devices and gadgets for data-driven storytelling. In *Data-driven storytelling*, pp. 135–149. AK Peters/CRC Press, 2018.
- [39] B. Lee, X. Hu, M. Cordeil, A. Prouzeau, B. Jenny, and T. Dwyer. Shared surfaces and spaces: Collaborative data visualisation in a co-located immersive environment. *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1171–1181, 2020.
- [40] J. Liu, A. Prouzeau, B. Ens, and T. Dwyer. Design and evaluation of interactive small multiples data visualisation in immersive spaces. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 588–597. IEEE, 2020.
- [41] W.-C. Lo, C.-L. Fan, J. Lee, C.-Y. Huang, K.-T. Chen, and C.-H. Hsu. 360 video viewing dataset in head-mounted virtual reality. In *Proceedings of ACM on Multimedia Systems Conference*, pp. 211–216, 2017.
- [42] MetaBallStudios. Waves and tsunamis comparison. <https://www.youtube.com/watch?v=2mK4SrjZwU8>, 2018.
- [43] MetaBallStudios. Ocean depth comparison. <https://www.youtube.com/watch?v=Q5C7sqVe2Vg>, 2021.
- [44] A. Prouzeau, M. Cordeil, C. Robin, B. Ens, B. H. Thomas, and T. Dwyer. Scaptics and highlight-planes: Immersive interaction techniques for finding occluded features in 3d scatterplots. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.
- [45] K. Rahimi, C. Banigan, and E. D. Ragan. Scene transitions and teleportation in virtual reality and the implications for spatial awareness and sickness. *IEEE Transactions on Visualization and Computer Graphics*, 26(6):2273–2287, 2018.
- [46] D. Ren, B. Lee, and T. Höllerer. Xrcreator: interactive construction of immersive data-driven stories. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2. ACM, 2018.
- [47] B. Saket, A. Endert, and J. Stasko. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pp. 133–142, 2016.
- [48] K. A. Satriadi, B. Ens, M. Cordeil, T. Czauderna, and B. Jenny. Maps around me: 3d multiview layouts in immersive spaces. *Proceedings of the ACM on Human-Computer Interaction*, 4(ISS):1–20, 2020.
- [49] A. Satyanarayan, D. Moritz, K. Wongsuphasawat, and J. Heer. Vegalite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350, 2016.
- [50] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.
- [51] D. Shi, F. Sun, X. Xu, X. Lan, D. Gotz, and N. Cao. Autoclips: An automatic approach to video generation from data facts. In *Computer Graphics Forum*, vol. 40, pp. 495–505. Wiley Online Library, 2021.
- [52] Y. Shi, X. Lan, J. Li, Z. Li, and N. Cao. Communicating with motion: A design space for animated visual narratives in data videos. In *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.
- [53] X. Shu, A. Wu, J. Tang, B. Bach, Y. Wu, and H. Qu. What makes a data-gif understandable? *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1492–1502, 2020.
- [54] StudioBinder. Guide to camera shots: Every shot size explained. <https://www.studiobinder.com/blog/types-of-camera-shots-sizes-in-film/>, 2020.
- [55] StudioBinder. Types of camera movements in film explained: Definitive guide. <https://www.studiobinder.com/blog/different-types-of-camera-movements-in-film>, 2020.
- [56] J. Tang, L. Yu, T. Tang, X. Shu, L. Ying, Y. Zhou, P. Ren, and Y. Wu. Narrative transitions in data videos. In *Proceedings of the IEEE Visualization Conference Short Papers*, 2020.
- [57] C. Tong, R. Roberts, R. Borgo, S. Walton, R. S. Laramée, K. Wegba, A. Lu, Y. Wang, H. Qu, Q. Luo, et al. Storytelling and visualization: An extended survey. *Information*, 9(3):65, 2018.
- [58] J. A. Wagner Filho, W. Stuerzlinger, and L. Nedel. Evaluating an immersive space-time cube geovisualization for intuitive trajectory data exploration. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):514–524, 2019.
- [59] M. Whitlock, S. Smart, and D. A. Szafrir. Graphical perception for immersive analytics. In *IEEE Conference on Virtual Reality and 3D User Interfaces*, pp. 616–625. IEEE, 2020.
- [60] A. Wu, Y. Wang, X. Shu, D. Moritz, W. Cui, H. Zhang, D. Zhang, and H. Qu. Ai4vis: Survey on artificial intelligence approaches for data visualization. *IEEE Transactions on Visualization and Computer Graphics*, 2021.
- [61] X. Xu, L. Yang, D. Yip, M. Fan, Z. Wei, and H. Qu. From wow to why: Guidelines for creating the opening of a data video with cinematic styles. *arXiv preprint arXiv:2202.02709*, 2022.
- [62] L. Yang, X. Xu, X. Lan, Z. Liu, S. Guo, Y. Shi, H. Qu, and N. Cao. A design space for applying the freytag’s pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):922–932, 2021.
- [63] Y. Yang, M. Cordeil, J. Beyer, T. Dwyer, K. Marriott, and H. Pfister. Embodied navigation in immersive abstract data visualization: Is overview+ detail or zooming better for 3d scatterplots? *IEEE Transactions on Visualization and Computer Graphics*, 27(2):1214–1224, 2020.
- [64] Y. Yang, T. Dwyer, B. Jenny, K. Marriott, M. Cordeil, and H. Chen. Origin-destination flow maps in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):693–703, 2018.
- [65] Y. Yang, T. Dwyer, K. Marriott, B. Jenny, and S. Goodwin. Tilt map: Interactive transitions between choropleth map, prism map and bar chart in immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(12):4507–4519, 2020.