Breaking the Fourth Wall of Data Stories through Interaction

Yang Shi, Tian Gao, Xiaohan Jiao, and Nan Cao

Abstract— Interaction is increasingly integrating into data stories to support data exploration and explanation. Interaction can also be combined with the narrative device, breaking the fourth wall (BTFW), to build a deeper connection between readers and data stories. BTFW interaction directly addresses readers by requiring their input. Such user input is then integrated into the narrative or visuals of data stories to encourage readers to inspect the stories more closely. In this work, we explore the design patterns of BTFW interaction commonly used in data stories. Six design patterns were identified through the analysis of 58 high-quality data stories collected from a range of online sources. Specifically, the data stories were categorized using a coding framework, including the input of BTFW interaction provided by readers and the output of BTFW interaction generated by data stories to respond to the input. To explore the benefits as well as concerns of using BTFW interaction, we conducted a three-session user study including the reading, interview, and recall sessions. The results of our user study suggested that BTFW interaction has a positive impact on self-story connection, user engagement, and information recall. We also discussed design implications to address the possible negative effects on the interactivity-comprehensibility balance, information privacy, and the learning curve of interaction brought by BTFW interaction.

Index Terms—Interaction, data-driven storytelling, narrative devices

1 Introduction

Data stories, as a communication medium, are increasingly integrating *interaction* to support exploring their essential elements: data, narrative, and visuals [30, 45, 67]. Various interaction techniques such as details-on-demand and hover highlighting [63] are leveraged by these data stories to improve the reading experience and facilitate data understanding. Interestingly, recent years have witnessed an emerging interactive approach to data stories, which requires input from readers. Notable examples include the Bloomberg article called "*Find out if your job will be automated*" [11] which allows a reader to search by his or her occupation and The New York Times article called "*You draw it*" [70] which enables a reader to draw his or her guess of a trendline on a chart. A key design idea in these stories is combining interaction with the narrative device, *breaking the fourth wall* (BTFW). Originating from theater, the fourth wall is an imaginary wall that separates actors from the audience, while the other three walls frame the scene [6]. To break the fourth wall, plays, television shows, and movies usually acknowledge the existence of the audience and speak to them directly. In data stories, the fourth wall convention also exists and can be flipped around through BTFW interaction, which directly addresses readers and asks for input related to themselves (e.g., type in *your* personal information, draw *your* expectation of data). Such user input is transformed into data and then incorporated into the narrative or visuals of stories to help establish a deeper connection between readers and stories [3].

In the visualization community, interaction for data-driven storytelling has received increasing interest. Prior research has sought to understand the effects of interaction designed to support data exploration and explanation. For example, McKenna et al. [49] compared data stories to be displayed with a step- and scroll-based navigation but did not find a significant difference in engagement between the two types of navigation input. Feng et al. [32] investigated the text-based search functionality of narrative visualization and found that this form of interaction can encourage user engagement and support information-seeking goals. Kim et al. [40] developed multiple elicitation techniques that incorporate users' prior knowledge in interaction (e.g., predict-only, predict-explain) and found that such interaction improves recall of data values. The aforementioned work explored specific interaction techniques and has laid a solid foundation for understanding them. However, the role of interaction that helps build reader connections to data stories has been largely overlooked.

To fill the gap, our work explores BTFW interaction in data stories and its common design patterns by following three complementary methods. First, we conducted a series of interviews with five data story experts to understand the benefits as well as challenges of applying BTFW interaction to data stories. Second, we collected a corpus of 58 data stories that leverage BTFW interaction, from a range of online sources. We then identified six design patterns of BTFW interaction, which were derived by coding i) the input of BTFW interaction provided by readers and ii) the output of BTFW interaction generated by data stories that responds to such input. Third, we conducted a user study with 109 participants to link user attitudes to those of the storytellers in the expert interviews. Qualitative feedback from the user study suggested that BTFW interaction has the potential to help build self-story connection, augment user engagement, and improve information recall. Also, concerns about the interactivity-comprehensibility balance, information privacy, and the learning curve of interaction were also raised when using BTFW interaction. The corpus, the design patterns, and the additional details of our user study can be accessed at our explorer of BTFW interaction, `https://idvxlab.com/btfwinteraction/`.

2 Related Work

To motivate our research, we review the literature on interaction for data visualization, interaction for data-driven storytelling, and breaking the fourth wall.

2.1 Interaction for Data Visualization

During the past decades, the ability of understanding data through visualizations has been augmented with the design and development of interaction techniques [29, 73]. To better understand the role of interaction in visualization, prior research has categorized interactions from different perspectives [34, 62, 77]. For example, Yi et al. [77] proposed seven general categories of interaction which were organized around users' intent when interacting with a visualization system, namely, Select, Explore, Reconfigure, Encode, Abstract/Elaborate, Filter, and Connect. Sedig and Parsons [62] synthesized 32 interactive actions in a framework of complex cognitive activities, and these actions were divided into two categories, including Unipolar and Bipolar.

Following-up studies [16, 59, 61] have further enriched the classification of interaction with the context information of interaction, which can bring the classification greater value in practice. For example, Roth [59] developed a taxonomy of interaction based on map visualization and deconstructed interaction into two primitives: objectives

• *Yang Shi, Tian Gao, Xiaohan Jiao, and Nan Cao are with Intelligent Big Data Visualization Lab at Tongji University.*
E-mail: {yangshi.idvx, gaotian, xh_xiaohan, nan.cao}@tongji.edu.cn.
Nan Cao is the corresponding author.

Manuscript received 31 March 2022; revised 1 July 2022; accepted 8 August 2022. Date of publication 26 September 2022; date of current version 2 December 2022. Digital Object Identifier no. 10.1109/TVCG.2022.3209409

which include user goals and interaction operands, and operators which involve enabling operators and work operators. Aiming to fill the gap of lacking distinction between the ends and means of a visualization task, Brehmer and Munzner [16] introduced a multi-level typology that can distinguish why and how a visualization task is performed, as well as what the task inputs and outputs are. The aforementioned work has greatly deepened the understanding of interaction, however, a fundamental question remains unclear: what is the consensual definition of interaction in data visualization? To answer this question, Dimara and Perin [29] synthesized an inclusive view of interaction and classified interactions into three categories: data actions, perceptualization data actions, and non-data actions. Our reading of the aforementioned work provides an initial understanding of the taxonomies of interaction for data visualization. Specifically, our work focuses on interaction for data-driven storytelling rather than visual analysis.

2.2 Interaction for Data-driven Storytelling

Data-driven storytelling is an increasingly adopted method that uses visualization as a storytelling form to communicate data insights [45]. Early in the development of data stories, researchers have noticed the important role that interaction can play. When Segel and Heer [63] first proposed the concept of narrative visualization, interactivity was set as a fundamental dimension of the design space. This dimension includes seven common interactions such as hover, filtering, and navigation. Their design space of narrative visualization has also been reflected and extended in recent years due to the rapid development of interaction. For example, Stolper et al. [66] identified and described storytelling techniques that were used in newly emerged data stories. These techniques were classified into four categories, including communicating narrative and explaining data, linking separated story elements, enhancing structure and navigation, and providing controlled exploration. Wang et al. [73] developed a design space of interaction for data comics, which describes interactions by a trigger, results in visual effects that can affect panel content, and/or panel layout around a particular communication goal.

Interaction can also provide direct benefits to story readers, including deepening the comprehension and recall of information, eliciting deeper reflection, and augmenting levels of engagement [40, 58, 78]. Specifically, Hohman et al. [36] identified five unique affordances of interactive articles, namely, connecting people and data, making systems playful, prompting self-reflection, personalizing reading, and reducing cognitive load. Some researchers have looked into the effects of a specific interaction technique on the reading experience of data stories [32, 35]. For example, Feng et al. [32] found that the search functionality can influence readers' information-seeking goals and alter the way readers engage with data. Compared with interaction for data visualization, interaction for data-driven storytelling is somehow in its infancy. To the best of our knowledge, no works have systematically explored the design space of interaction for data-driven storytelling. Also, the role and effects of many widely used interaction techniques have not been fully understood yet. Thus, our work attempts to understand the interaction that is used to break the fourth wall of data stories and explore the possible benefits of such interaction.

2.3 Breaking the Forth Wall

The fourth wall is a notion first proposed by Denis Diderot in the 18th-century theater [9, 17, 26]. In theatrical performance, the fourth wall is an imagined, invisible wall between actors and the audience. While the audience can see through this wall, the actors cannot, thus, are not aware of the existence of the audience [72, 75]. Nowadays, the fourth wall has become a well-established convention followed by various mediums, such as theater, cinema, and literature.

Breaking the fourth wall (BTFW) is when the fourth wall convention is violated in performance. To achieve the goal of BTFW, various strategies have been adopted and they can be divided into two categories based on mediums. For mediums such as films and television shows, BTFW usually takes the form of a one-way expression by the character to the audience, due to the separation of the performance recording and spectating [19]. Specifically, the two most common strategies include directly addressing the audience [5, 21, 72] and using camera techniques [50, 51] such as having actors look at the camera straightly to establish eye contact with the audience and utilizing close-ups to promote a sense of physical proximity. In literature, a frequently used strategy is metafiction, in which characters show their awareness that they are fictitious being [74]. A well-known example of BTFW is Jane Eyre [18] and the most famous line in the novel says, "*reader, I married him.*" The other category involves mediums such as live theater and video games, BTFW often takes the form of back-and-forth interactions between characters and the audience/players [9, 22, 25]. For both categories, using BTFW is believed to help create a entertaining atmosphere [4], strengthen involvement [5, 13], and elicit empathy [19].

As an established and accepted theory, BTFW has been drawn on in many fields (e.g., archaeology [27], education [68], psychology [56]). However, it has received limited attention in the community of visualization. In one case, Bach et al. [6] suggested that BTFW can be used as a narrative design pattern to frame data-driven storytelling and trigger affective responses. Instead, we investigated how to achieve BTFW in data stories by directly "speaking" to readers through interaction. To the best of our knowledge, our work systematically explores the combination of BTFW interaction and data-driven storytelling for the first time.

3 Preliminary Study

To understand how BTFW interaction has been leveraged by storytellers in practice, we conducted a series of semi-structured interviews with five domain experts. Our goal is to gain insights into the following issues: (i) the motivations for applying BTFW interaction to data stories and (ii) the potential challenges arising from using BTFW interaction.

3.1 Interviews with Domain Experts

We recruited domain experts by publishing an invitation poster on social media platforms and word of mouth. To understand perspectives from both academia and industry, we finally invited three data journalists E1 (six years' experience), E2 (six years' experience), E3 (seven years' experience) from news agencies, and two data visualization researchers E4 (seven years' experience), E5 (six years' experience) from research labs. We introduced the concept of BTFW interaction to the experts and all of them reported that they are experienced in applying it to designing data stories. In their past practice, the experts sometimes collected basic personal information for their data stories such as age and residential address. The experts considered the collection of data as a means of helping create personalized stories. They decided to use such interaction because, for certain topics, the content of the story can vary from person to person, and a generalized description will reduce the degree of personalization. "*For stories that focus on personal health, I need the user's information to rate and advise him on his health status*" (E2). In addition, they applied such interaction to show the gap between facts and public perception. "*It reminds readers that their perception of a good body may be opposed to the standard*" (E3).

Each interview was conducted online and included two sessions. In the first session, we introduced our research topic and the concept of BTFW interaction to the interviewees, We also showed them examples of data stories that use BTFW interaction and encouraged them to share such data stories created by their own. After the interviewees indicated that they were clear about the topic and the concept, we started the second session. We asked them a set of interview questions such as "*what type of BTFW interaction did you use in your data story*?", "*why did you add BTFW interaction to your data story*?", and "*what challenges did you face when using BTFW interaction*?". Each interview lasted about 40 minutes and was audio-recorded for subsequent analysis.

3.2 Analysis and Findings

To analyze the qualitative data collected from the interviews, we transcribed the audio recordings and then coded the data based on our two research questions following the thematic analysis process [15]. Specifically, two researchers first independently read through all the transcriptions and highlighted the comments related to the questions.

Next, we independently generated codes from the highlighted comments and then grouped similar codes until reaching a saturation point. Finally, we met for two sessions to compare our codes, group similar codes, and discuss mismatches until we reached a 100% agreement on the final themes.

Motivations. Three themes emerged as possible benefits of applying BTFW interaction to data stories.

Build a self-story connection. All of the participants mentioned that applying BTFW interaction is an effective method to build a connection between a reader and a data story. This benefit arises because BTFW interaction offers a way to "put yourself in the story". For example, E3 said that "*when discussing serious topics such as well-being and environment, it can be difficult for readers to relate themselves to the topics. However, with this type of interaction, readers' data can be integrated into the narrative and they'll feel more connected to the stories*". E5 agreed that "*through this interaction, the connection between the content of a story and a reader's personal experiences is strengthened and the reader's identification with the core message of the story can be evoked*". E4 complemented that "*using this interaction feels like I'm creating my own data story*".

Augment user engagement. Another theme mentioned by all of the participants is that BTFW interaction serves as a storytelling device for attracting attention and augmenting engagement. Both E2 and E5 mentioned that "*the inclusion of such interaction can arouse readers' interest, elicit their emotions, trigger their desire to share and discuss, and even induce their willingness to act.*". Similarly, E4 stated that "*statistics can be bored or even cold sometimes, adding interaction to break the fourth wall helps interest the reader in consuming the data and enticing them to inspect it more closely*". On the other hand, E1, E2, and E3 explained this benefit in terms of "*playfulness*" provided by interaction. Specifically, E1 said "*compared to other types of interaction, it allows for more initiative; it encourages readers to try and play and motivates them to explore the data in greater depth*".

Improve information recall. Three participants (E1, E3, E5) mentioned that applying BTFW interaction can improve information recall. For example, E1 stated that "*information related to personal data is more likely to stick in memory*". E5 said that "*compared to simply browsing text, readers become more active when interacting with data and visualization, which can leave a lasting impression*". E3 complemented that "*inputting individual facts and opinions can create a sense of participation. Such information can be easily recalled afterward*".

Challenges Three themes emerged as the challenges of applying BTFW interaction to data stories.

Information privacy concerns. Four participants (E1, E2, E3, E4) mentioned that readers' concerns about personal privacy may reduce their willingness to use BTFW interaction. E2 commented that "*typing in personal data feels like self-disclosure for readers, and they are likely to be concerned about whether this action will lead to any additional consequences beyond contributing to story reading*". E3 complemented that "*for some readers, data such as weight, age, and gender are quite sensitive and they are not always willing to disclose them. To protect their privacy, they may simply give up reading, or may enter false data to test the result*". E1 added that "*storytellers have an obligation to inform and explain how the information entered by readers is going to be used*".

The balance between interactivity and comprehensibility. Three participants (E1, E2, E5) mentioned that using BTFW interaction may disrupt the balance between the interactivity and comprehensibility of data stories: "*readers may not always have a strong desire to explore and interpret on their own, especially for simple topics. They may prefer to read concise and 'easy to digest' stories; unnecessary interactions can be a barrier to effective reading*"(E2), "*the key task in crafting a data story is conveying its core idea determined by its author. However, if a reader is immersed in the world of self that was found in the story, the author's core idea may be ignored*"(E5).

The learning curve of interaction Four experts (E2, E3, E4, E5) noted that readers may be intimidated by the steep learning curve of interaction. "*The mechanisms of this interaction can be complex sometimes, and if readers have difficulty figuring it out in a short time they will simply skip it. It's even more challenging when reading on mobile devices becomes more prevalent*", said E5. E2 complemented, "*learning how to use this interaction will increase the cognitive load of readers compared to more basic and common interactions like hover and scroll; this is also the cost of novelty and playfulness*". E3 added that "*when applying such interactions to my stories, I will try to minimize the learning cost for readers by giving them clear guidelines or deconstructing the interaction into multiple sub-steps*".

4 Coding Framework and Design Patterns

Interaction can be used to break the fourth wall of data stories by directly addressing readers and requiring their input concerning themselves. To explore BTFW interaction, we first collected and analyzed a corpus of 58 data stories. Then, we derived a set of design patterns of BTFW interaction from the coding of our corpus.

4.1 Methodology

We analyzed BTFW interaction in a two-step methodology. First, we collected a corpus of data stories that use BTFW interaction. Then, we constructed a coding framework that integrates interdisciplinary knowledge to identify BTFW interaction patterns.

4.1.1 Data Collection

To collect a corpus of high-quality data stories leveraging BTFW interaction, we started with the data story corpora generated by previous studies [36, 43, 49, 63, 66, 67]. After merging these corpora and removing duplicates, we obtained 358 data stories and successfully retrieved 250 of them from the internet. These data stories were collected from well-known sources (e.g., New York Times, Washington Post), recommended by domain experts (e.g., experienced data journalists and visualization designers), or popular among readers (e.g., winners of the Sigma Awards, opinionated lists of best visualizations). Next, we filtered out 160 data stories that are non-interactive, which use only scrolling and stepping for navigation [49]. Following the methodology by [2, 44, 64], we then searched for additional samples by surveying news agencies famous for data stories (e.g., The New York Times, FiveThirtyEight) and authoritative data visualization awards (e.g., Information is Beautiful, The Sigma Award). For example, we first located related columns on these platforms such as *The Upshot* in The New York Times and then searched for news articles with keywords such as "interactive article" and "interactive infographic". After that, we complemented 176 additional interactive data stories. Finally, we checked

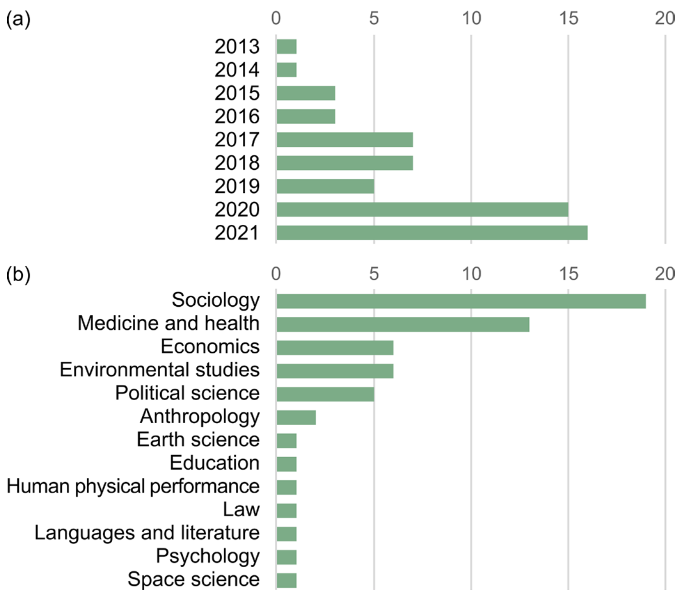

Fig. 1. Frequency of the 58 data stories about (a) year and (b) subject.

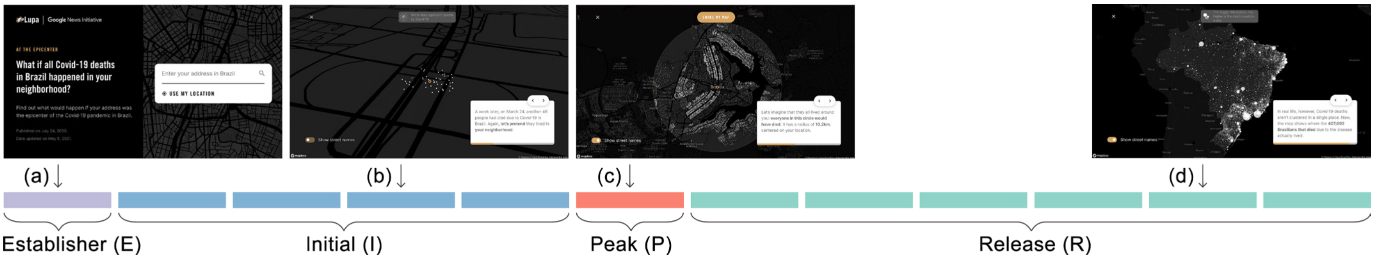

Fig. 2. Coding of an example data story entitled *What if all Covid-19 deaths in Brazil happened in your neighborhood* [48]. Screenshots from the story are associated with the corresponding narrative category [2]. The description of each screenshot says, (a) "*Enter your address in Brazil*", (b) "*A week later, on March 24, another 46 people had died due to Covid-19 in Brazil. Again, let's pretend they lived in your neighborhood.*" (c) "*Let's imagine that they all lived around you: everyone in this circle would have died. It has a radius of 10.2km, centered on your location.*" and (d) "*In real life, Covid-19 deaths aren't clustered in a single place. Now, the map shows where the 407,092 Brazilians that died due to the disease actually lived.*"

these interactive data stories to ensure that each of them is data-driven and uses BTFW interaction. Specifically, we set three criteria to mandate that the stories (i) present arguments or facts supported by data, (ii) include at least one data visualization, and (iii) address readers directly and require their input. As a result, we identified 58 out of 266 data stories to constitute our final corpus. The 58 data stories were produced from 2013 to 2021 and their subjects range from sociology to space science, as shown in Fig. 1.

4.1.2 Coding Framework

After data collection, we coded the corpus from two aspects using the method of close reading [20]: (i) what is the *input* of BTFW interaction and (ii) what is the *output* of BTFW interaction that responds to the input. The two-dimensional framework mirrors a key concept of interaction in Human-Computer Interaction (HCI) [37] and visualization literature [29], which considers interaction as *dialogue*. By using the metaphor of dialogue, interaction is viewed as a cycle of communication between a user and an interactive system, in which actions from the user and reactions from the system can be both perceived [29]. Similarly, our coding framework models BTFW interaction as a dialogue between readers and interactive data stories. In such a dialogue, we considered the actions from readers as the input of BTFW interaction while the reactions from stories as the output.

For (i), we analyzed the input of BTFW interaction (the actions from readers) through the lens of the data–information–knowledge–wisdom (DIKW) hierarchy [1, 60], which is a widely recognized theory that originated from the field of information management. Specifically, the hierarchy explains how the human mind moves data up to higher levels by progressive organization; first comes data, next are information and knowledge, and finally comes wisdom. We think that the input of BTFW interaction from readers also goes up in the higher levels of the hierarchy, from data to wisdom. Thus, we categorized the input of BTFW interaction as follows,

- **Data**. Data is a recording of a fact situated in a specific context. It is often unorganized and unprocessed. For example, "*input your address to see the medical condition nearby*" is input of data as it simply asks for a geographic location, which is a fact.
- **Information**. Information is data processed and organized for a purpose, adding value to the understanding of a subject. For example, "*how do you protect yourself from Covid-19?*" asks input of information as readers need to review their daily routine and then choose the answer from options such as "wear a mask", "get vaccinated", and "avoid crowds".
- **Knowledge**. Knowledge is information merged with expert experience and skills, resulting in a valuable asset that can be used to support decision-making. For example, "*draw your prediction of the epidemic trend of Covid-19 in New York*" asks input of knowledge as readers make predictions by reasoning about the information of historical trends.
- **Wisdom**. Wisdom is the accumulation of knowledge, allowing people to apply what they learned from one domain to solve new problems. "*Develop a reasonable and effective plan to stop the spread of Covid-19 in New York*" requires the input of wisdom as readers need to synthesize their knowledge such as medicine and public administration to provide a solution.

For (ii), we coded the output of BTFW interaction (the reactions from stories) by analyzing how it integrates the input into the narrative or visuals of stories. To do this, we used the narrative structure developed by Amini et al. [2]. This narrative structure, based on Freytag's Pyramid [33] and Cohn's narrative grammar [23], is used to classify the sequences of data videos regarding their role in the narrative, including Establisher, Initial, Peak, and Release (EIPR). Inspired by the EIPR structure, we classified the output of BTFW interaction as follows,

- **Establisher**. Establisher sets the context for a data story to introduce the topic and raise interest. For example, in "*What if all Covid-19 deaths in Brazil happened in your neighborhood?*" [48], the Establisher displayed in Fig. 2 (a) sets up the data story by requiring readers to "*enter your address in Brazil*", attracting their attention and quickly transitioning to the Initial.

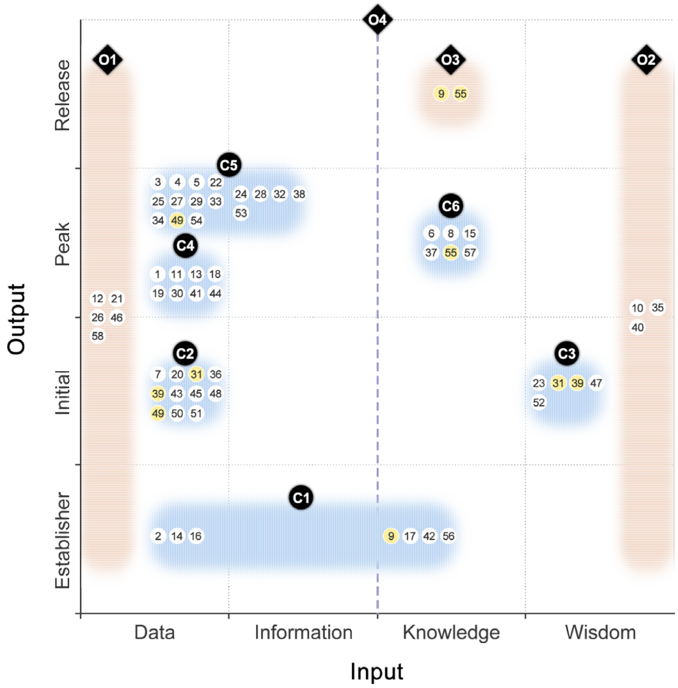

Fig. 3. Our framework to code BTFW interaction in the 58 data stories in our corpus. Input (DIKW) is represented as the x-axis while output (EIPR) is represented as the y-axis. Each data story is illustrated as a dot. Dots are grouped into a cluster as they share common characteristics and thus form a design pattern. The details of each data story can be found in our explorer, `https://idvxlab.com/btfwinteraction/`

- **Initial**. Initial lays the foundation for the upcoming tension of the story. Fig. 2 (b) shows the Initial which illustrates the increasing number of deaths by Covid-19, leading to the Peak.
- **Peak**. Peak reveals the climax of tension, that is, the major insights of the story. Fig. 2 (c) shows the Peak, where the major insight, the number of deaths nationwide, is revealed. The number is visualized as a circle on the map centered on the readers' location, answering the question asked at the beginning of the story.
- **Release**. Release shows the aftermath of the Peak, providing additional information around the major insight. Fig. 2 (d) illustrates the Release which complements additional information about the Peak such as the distribution of deaths in the whole country.

Based on the coding framework, two researchers independently coded the corpus. We met for three sessions to compare our codes and discuss mismatches until we reached a 100% consensus. Fig. 3 shows the final codes of the samples in our corpus. Specifically, the x-axis represents input (DIKW) while the y-axis represents output (EIPR). Each sample in our corpus is visualized as a dot with a unique ID and is placed in the grid according to its input and output. A sample is colored in yellow if it uses multiple BTFW interaction techniques and can be classified into more than one category. After coding, we found that these samples can be grouped into clusters (colored in blue), where samples share similar characteristics and constitute a design pattern. Based on the observation, we identified six design patterns of BTFW interaction (**C1**-**C6**) that are commonly used in data stories. The next subsection explains these design patterns in detail and relates them back to the samples in our corpus.

4.2 Design Patterns

A design pattern fills one cluster in Fig. 3 and is further defined by a name, a description, an example, and an abstract illustration, as shown in Fig. 4. In the following, we describe the identified design pattern from **C1** to **C6** (left-right, down-top in Fig. 3).

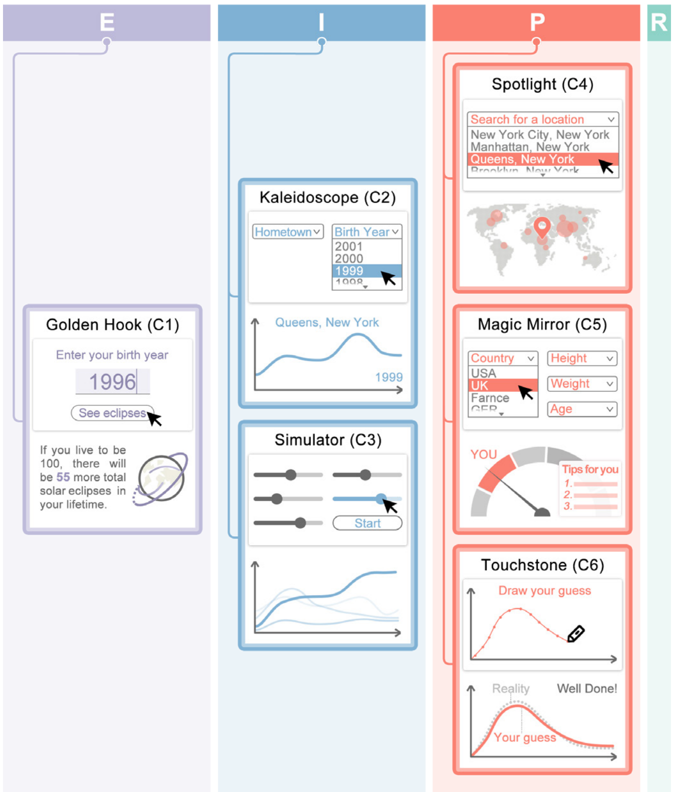

Fig. 4. The design patterns of BTFW interaction, marked as **C1**-**C6**. The output of BTFW interaction is labeled as E (Establisher), I (Initial), P (Peak), and R (Release) on the left of each design pattern.

Golden Hook (C1). Using Golden Hook, a data story is started with the Establisher based on readers' personal data or knowledge. This design pattern can only be found at the beginning of the story and is used to help readers quickly understand the topic of a story. Much like a fish gets hooked by bait, Golden Hook attempts to captivate readers and make them want to read more. For example, at the Establisher of "*Here's every total solar eclipse happening in your lifetime*" [46], the story asks for readers' birth year, then calculates the number of total solar eclipses in readers' lifetime and visualizes the path of each eclipse. Although the results may be surprising to readers, the story does not explain the reason but goes on to introduce the formation principle of total solar eclipse. Although the visualization of user input is no longer referred to at the Initial or Peak of the story, it helps reveal unexpected findings and increases curiosity from readers. Here, we merged inputs of data and knowledge, as the data stories using these two levels of input share similar characteristics in forms and serve a specific purpose, that is, attracting readers' attention quickly and motivating them to find explanations in the following story.

Kaleidoscope (C2). Kaleidoscope is used at the Initial of a data story, where supporting facts are provided to reveal deeper insights. This design pattern requires readers' personal data as a supplement to the supporting facts and then presents the data as personalized visualization. It serves the intent of persuading by delivering more convincing arguments backed with data provided by readers. Different from Golden Hook (C1) that visualizes user input to grab their attention, Kaleidoscope uses the data from readers directly to help reveal the core message. In this way, readers are effectively engaged in the process of reasoning and are more likely to support the arguments presented by the storytellers. For example, "*Why we're experiencing so many unusually hot summer nights*" [10] first introduces that people are experiencing unusually hot summer nights and then encourages readers to look up their cities. Next, the temperature distribution of the summer night in the reader's city is visualized and presented. by observing the visualization, the reader now understands that in his city, more nights are abnormally hot, which foreshadows the major insight that will be explained in the Peak: global warming.

Simulator (C3). Simulator uses visualization as the input interface for exploration in the Initial of a data story. Specifically, the visualization models a fact or a theory and exposes parameters that readers can manipulate to change the behavior of the simulation. By making a series of decisions, readers can learn the fact or the theory by "experiencing" it themselves. In contrast to Kaleidoscope (C2), Simulator strengthens arguments by encouraging readers to use trial and error rather than directly presenting the outcome to them. In this way, this pattern engages readers in the exploration and discovery of the arguments that are not explicated by the storytellers. For example, "*Should prison sentences be based on crimes that haven't been committed yet?*" [7] first explains the dilemma of awarding parole. Then, it instructs readers to set cutoffs for the risk categories of offenders. After multiple attempts, readers learn through "play" that even the best risk assessments yield probabilities, not certainties; parole-eligible prisoners who are mistakenly labeled as "low risk" can be awarded but probably reoffend in the future.

Spotlight (C4). Spotlight is applied at the Peak of a data story, where the major insight is shared with readers. With this pattern, the major insight is presented as a visualization, usually in a form of a map, to encourage readers to explore. Specifically, the visualization asks for readers' addresses and highlights the data points related to them. Compared to Kaleidoscope (C2) which uses personalized visualization, Spotlight provides an overview first and then allows readers to zoom in to obtain a closer view of interest. Moreover, Spotlight is commonly used in data stories around serious topics such as well-being and environment to support readers in relating individual lives to society. For example, to introduce the opioid crisis in America, "*How many pain pills went to your pharmacy?*" [31] presents a map to show the number of opioid pills that were handled by pharmacies across the country. On the map, readers can search for their address to find how many pills were shipped to nearby pharmacies, acquiring a deeper understanding that they are in the midst of the opioid crisis.

Magic Mirror (C5). Magic Mirror is used at the Peak of a data story and can produce the most personalized story among all the design patterns. Stories combined with this pattern usually take the form of a quiz or a calculator, integrating readers' data or information into the communication of a highly customized major insight. Different from Spotlight (C4) which focuses on series topics, Magic Mirror is often used to support discussion on topics about individual lives such as health and housing. This pattern helps reflect "who you are" and more importantly, provides personalized advice to improve one's current situation. For example, "*BMI and obesity: where are you on the UK fat scale?*" [8] first introduces that the majority of adults are overweight in the UK. Then, it requires readers' personal data such as weight and age to calculate one's body mass index (BMI). In the calculated result, in addition to their BMI, readers learn how they compared with the rest of the nation and get tips from health experts on how to improve their health conditions.

Touchstone (C6) Touchstone helps reveal the major insight in the Peak of a data story by prompting readers to make a guess. Before showing the actual data, Touchstone asks readers to predict it based on their knowledge. Then, readers are shown the visualization of the actual data against their prediction. Meanwhile, visual and textual annotations are added to the visualization to emphasize the difference between readers' guesses and the actual data. In this way, readers are more likely to reflect on the gap and be impressed with the major insight. This design pattern serves as a "touchstone", by which the understanding of a fact or a concept can be tested. For example, "*You draw it: how family income predicts children's college chances*" [70] first raises a question about how likely is it that children who grow up in poor and rich families go to college. Then, it asks readers to draw their guess for this question on a blank chart. After that, the reality is revealed and compared with readers' answers.

4.3 Design Patterns Observations

In addition to the six design patterns, we also derived two additional findings by observing the distribution of the samples in our corpus, marked as **O1**-**O4** (colored in pink) in Fig. 3.

The first finding is about the two vertical clusters, O1 and O2. Data stories in O1 and O2 use BTFW interaction in all the narrative categories of the EIPR structures while data stories in C1-C6 integrate BTFW interaction into one narrative category. Specifically, O1 usually requires readers to input personal data at the very beginning and provide highly personalized content throughout the stories. For example, "*What if all Covid-19 deaths in Brazil happened in your neighborhood?*" [48] (Fig. 2) is an exemplar of O1, which describes the pandemic related to individuals from the beginning to the end. When compared to C1, C2, C4, and C5, O1 offers a more immersive reading experience. However, due to its high level of personalization, O1 is limited in relying heavily on narrative messaging to provide observations and explanations about insights. Similarly, stories in O2 are highly personalized but based on readers' input of wisdom rather than data. These stories usually take the form of news games [12], which construct simulation models where readers can manipulate parameters to make choices throughout the stories. Readers' decisions can have a significant impact on the current narrative as well as the ending. For example, in the news game, "People of the pandemic" [53], readers are required to make a series of decisions on how frequently people leave their homes each week during the pandemic while the final goal is to save as many lives as possible. With news games, storytellers can bring entertainment to readers and communicate insights through "try and play" similar to C3 and C6. On the other side of the coin, O2 is usually limited in allowing core messages to be explicated due to its high interactivity.

In addition to O1 and O2, stories in O3 also use BTFW interaction in more than one narrative category of the EIPR structure. Similar to C1 and C6, O3 generates personalized narrative or visualization based on user input of knowledge. However, O3 also reveals how readers perform when compared to other people in the Release, where solution and next steps are usually recommended. For example, "*You draw it: how family income predicts children's college chances*" (C6) [70] also shows if one's guess is more or less common when compared to others at the end. Such social comparison can result in a better assessment of oneself and thus promote a deeper self-reflection. Given its characteristics, O3 can be used to "callback" the input previously required by the story to surprise readers.

The second finding is about the central axis (O4) of our framework, as shown in Fig. 3. When analyzing the clusters on the right of the central axis, we found that they are characterized by establishing criteria for evaluating user input (Knowledge and Wisdom). For example, C3 estimates the behaviors of the simulation by telling readers if they achieve the goal set by the storytellers. C6 evaluates readers' guesses by showing the gap between their prediction and reality. O2 assesses readers' decision-making by showing if a story has a "happy ending" while O3 assesses readers' knowledge about a fact or a theory by comparing it to that of other people. The reason is that all these patterns tend to promote self-explanation and self-reflection among readers. On the other hand, clusters on the left of the central axis focus on visualizing instead of evaluating user input (Data and Information). They are characterized by providing readers with personalized data facts, resulting in a better connection between readers and stories.

5 User Study

After identifying the six design patterns of BTFW interaction, we conducted a user study to understand its benefits and concerns. The goal is to link user attitudes to those of the storytellers in the expert interviews and see if these were aligned. Specifically, we were interested in (i) the benefits of using BTFW interaction, including *self-story connection, user engagement*, and *information recall*, and (ii) the concerns of using BTFW interaction, including *information privacy, the balance between interactivity and comprehensibility*, and *the learning curve of interaction*. Our study mainly focused on collecting and analyzing the qualitative responses from the participants while the quantitative results are available online, `https://idvxlab.com/btfwinteraction/`.

5.1 Stimuli

In the user study, we compare data stories leveraging BTFW interaction with a non-interactive data story, as shown in Fig. 5. Specifically, six data stories were selected from our corpus, one for each of the design patterns of BTFW interaction (**C1-C6**). We also mandated that all of the data stories are on pandemic-related topics to provide situations where the participants are likely to have shared knowledge or expectations of the domain. For example, C1 asks for readers' personal data and suggests their position in the vaccine line at the beginning of the story. C4 requires readers' addresses to see the capacity of hospitals nearby them. C6 invites readers to guess how much prices have risen for common grocery items during the pandemic. We were unable to find a pandemic-related data story employing the Magic Mirror (C5) technique. As an alternative, we selected the data story, "*How you will die*" [76], whose health-related topic is the most related to the pandemic in our corpus. We also included one non-interactive data story about Covid-19 (**C0**). Although without interactivity, it breaks the fourth wall through the narrative by addressing readers directly such as "*the risk of encountering a COVID-19-infected person at YOUR small, intimate gathering was about 82 percent*". The seven data stories (**C0-C6**) constitute the stimuli for our study.

5.2 Participants

To estimate the appropriate sample size, we first performed a power analysis in one-way ANOVA based on our pilot study. We achieved 0.8 power under $\alpha = 0.05$ with 42 participants per condition, resulting in at least 108 participants in total. Then, we recruited 109 participants (55 females) by posting advertisements on social platforms. The participants ranged in age from 21 to 50 (M = 24.54, SD = 3.15). The majority of them received Master's degrees or above (Bachelor's degree: 25 (22.9%), Master's degree or above: 84 (77.1%)) and varied in educational backgrounds (e.g., design, computer science, electrical engineering, business administration). The breakdown of their expertise in visualization is as follows: novice: 36 (33.0%), advanced beginner: 41 (37.6%), competent: 17 (15.6%), proficient: 11 (10.1%), expert:

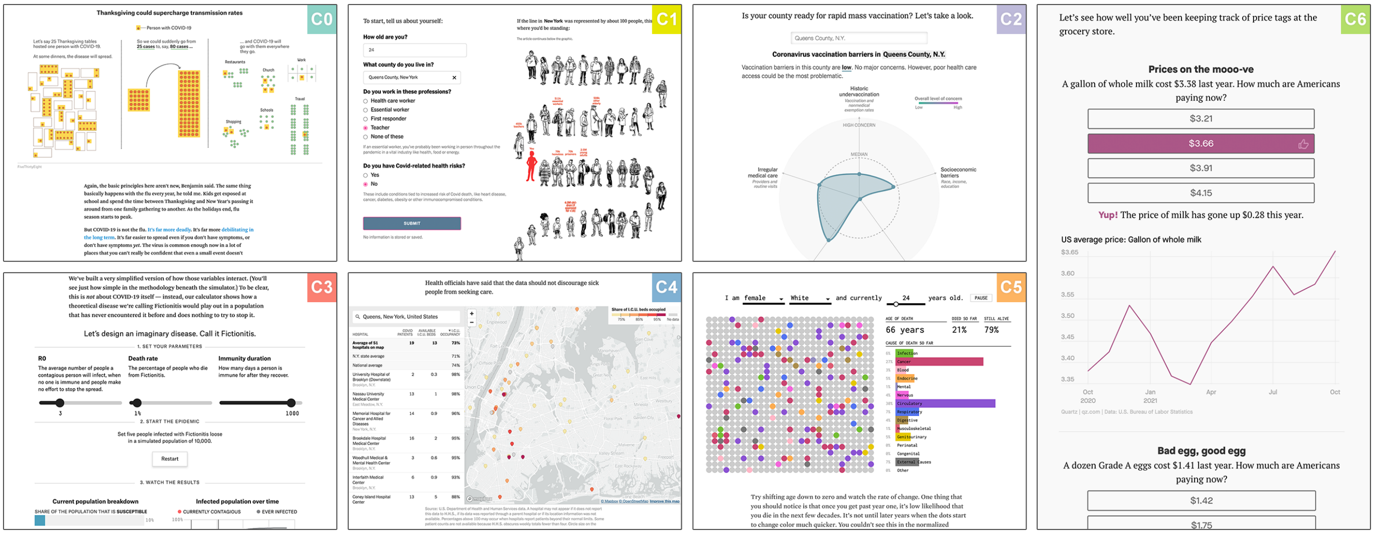

Fig. 5. Seven data stories about Covid-19 collected for C0 (*Why even a small thanksgiving is dangerous* [41]), C1 (*Find your place in the vaccine line* [69]), C2 (*What are the vaccine roadblocks where you live?* [38]), C3 (*Without a vaccine, herd immunity won't save us* [57], C4 (*How full are hospital I.C.U.s near you?* [24]), C5 (*How you will die* [76]), and C6 (*Guess how much prices on 11 common grocery items have risen in one year* [28]), and used as the stimuli in our user study.

4 (3.7%). The participants also have different cultural backgrounds, including China: 84 (77.1%) and the United States: 25 (22.9%).

5.3 Task and Procedure

The user study consists of three sessions: the *reading* session, the *interview* session, and the *recall* session. Before the study, we sent emails to our participants to ask about their willingness to take part in the interview session. As a result, 26 out of the 109 participants responded positively. In the reading session, we started with a 10-minute introduction explaining the goal of our study to the participants and obtaining their consent to audio recording. Also, the participants were instructed to provide their demographic information such as age, gender, and educational background. Then, they were presented with two interactive data stories chosen from our stimuli and the non-interactive data story (C0), one at a time. We ensured that each of C1-C6 was read by at least 36 participants. The order of the three stories was randomized to avoid confounding effects. After the participants finished reading and exploring a story, they were asked to rate it using a questionnaire that measures the design patterns from four aspects, including *self-story connection, user engagement, information recall*, and *the balance between interactivity and comprehensibility*. When all three data stories had been read, we started the interview session. We conducted a semi-structured interview with the participants to learn about their reading experiences and the reasons for ratings. The reading and interview sessions lasted about 45 minutes and 30 minutes, respectively.

Two weeks after the interview session, we sent emails again to invite our participants to complete the recall session. 18 participants accepted our invitation. At the beginning of the recall session, to activate their memory about the three stories they had read, the participants were asked to retell the details of each story as much as possible. After that, we asked several questions like "Do you remember any specific data in these stories?", "What is the input and output of BTFW interaction?", and "What are the insights revealed in these stories?". Each participant finished the recall task in about 15 minutes.

6 Results and Analysis

After the user study, we collected the quantitative data from the 109 questionnaires and transcribed the qualitative data from the interviews and the recall session with the participants. The qualitative data were analyzed using thematic analysis and seven themes emerged. In the following, we described the seven themes to understand the participants' thoughts, comments, and suggestions.

Self-Story Connection and Engagement Based on the quantitative analysis, we can observe that using BTFW interaction in data stories results in stronger *self-story connection* and *engagement* among readers when compared to C0. With the help of the personalized visualization and narrative, the participants could find the connection between the story and themselves and thus become more engaged in reading (e.g., "*I felt that C4 is quite close to my daily life, as its narrative was organized around my address.*" (P13), "*When I realized that the following narrative could lead me to understand my personalized visualization in-depth, I felt very curious and was eager to read it.*" (P7), "*By breaking the fourth wall, you create more intimacy between the storytellers and the audience, and it helps strengthen their relationship.*" (P2)). In addition, we observe that user input from different categories of the DIKW hierarchy appeals to the participants in different ways. For data and information, address-related input can especially engage readers in the story (e.g., "*compared to C5, which asked me to enter my gender and age, C2 asked me to enter my address, which made me feel more relevant to the story*" (P9)). When asked about the input of knowledge, the participants found it brings enjoyment as it usually comes up with the form of examination (e.g., "*I was surprised that I could get seven questions right in C6, it gave me a strong sense of accomplishment*" (P4), "*Guessing the price was like playing a game and I was always looking forward to the next level*" (P2)). For stories that require the input of wisdom, many participants attributed the sense of connection primarily to their in-depth thinking when attempting to follow the storyteller's core idea. For example, P27 noted, "*while adjusting the parameters in C3, I referred to my daily experiences and carefully decided the value for each parameter. I was highly immersed in this process as if I were controlling the world in the story*".

Information Recall In general, the participants' memory of the stories faded a lot after two weeks. The topic of interactive data stories left a deeper impression on the participants than the non-interactive story. Most of the participants remembered the topic of C1-C6 clearly (e.g., "*it's about the cause of death in different age groups*" (P19)). By contrast, 27.8% of the participants almost completely forgot C0 (e.g., "*I thought I only read two stories*" (P12)). For the participants who still remembered C0, they could also retell the major insight of it (e.g., "*it told us small gathering is also dangerous during the pandemic*" (P18)). In terms of C0-C6, what the participants remembered was not only the major insight. Many participants mentioned several minor points that they gained from their observations of the interactions (e.g., "*raising the value of R0 is the most effective way to get a sharp curve of infected* *population*" (P1)). Besides, in most cases, the participants remembered the required input of BTFW interaction clearly (e.g., "*I could adjust the value of R0, the death rate, and immunity duration of a fiction virus*" (P4)), as well as the form of the output (e.g., "*it is a map visualizing the hospital capacity*" (P17)). We also observed that none of the participants remembered specific data values in the stories, but many of them could roughly retell the result of the interaction (e.g., "*I find my place in the vaccine line a little forward from the middle*" (P16)).

Story Preference When asked about which story was their favorite, most of the participants chose stories with BTFW interaction and their reasons vary, including learning important facts (e.g., "*knowing the capacity of the hospital around me may save my life in the emergency*" (P8)), feeling immersed into the narrative (e.g., "*I was absorbed in the plot and empathized with the storyteller*" (P2)), and having great enjoyment (e.g., "*the interaction reminds me one of my favorite video games*" (P1)). However, some participants mentioned that interactive data stories were not their favorites because they sometimes felt confused during interaction (e.g., "*the authors do not provide guidance clear enough, I can learn nothing from the story*" (P16)). This feedback is in line with our findings from the preliminary study that when inappropriately designed, interaction may hinder comprehensibility. In addition, the participants preferred interactive stories as the stimuli that were most relevant to themselves. The sense of relevance is mainly from the personalized narrative or visualization (e.g., "*C0 is for everyone, but I feel that the comic of the vaccine line in C1 was exclusively designed for me*" (P25)). More than half of the participants chose the story most related to themselves as their favorite story. Also, in 30.8% of the cases, C0 was chosen as the most related story. These participants explained that they chose C0 mainly because of its topic (e.g., "*small gathering is one of the most common daily activities in my life, it happens much more frequently than vaccination*" (P15)).

The Role of BTFW Interaction We also found that most of the participants roughly classified the role of BTFW interaction into three narrative categories, namely beginning, middle, and ending (e.g., "*all these stories followed the typical three-phase syllogism, and I felt the interaction made different contributions to each phase*" (P26)). When asked about which narrative category was the most ideal one for integrating BTFW interaction, all of the three categories were mentioned by the participants. 10 out of 26 participants valued an inspiring beginning most (e.g., "*readers' interest should be raised as early as possible*" (P3)), some believed BTFW interaction can be most effective only with an adequate introduction and explanation (e.g., "*I could not fully understand the interaction without a clear context*" (P12), "*I wish to interact with the climax of the story and see my personal data turns out to be*" (P10)), while a few participants mentioned they hoped to experience BTFW interaction at the end of a data story (e.g., "*I could make the most use of the interaction after I learned what's going on*" (P18), "*if the storyteller could integrate his conclusion with the visualization result of my data produced in former stages, I could have a deeper understanding of the story*" (P24)).

Information Privacy and Reading Experience For most of the participants, in order to get personalized stories, the potential privacy risk of BTFW interaction was acceptable (e.g., "*It sounds fair to me if I can make this story really relevant to myself by providing some less critical personal information*" (P7)). On the other hand, for several participants, the willingness to interact decreased due to their concerns about information privacy. Specifically, the degree of concern is related to the type of required input. For example, the participants were more sensitive if data stories require data and information as input, such as age (e.g., "*It feels like being asked how old I am by strangers in real life and it's offensive to me.*" (P12)) and address (e.g., "*Address is the critical information for my personal security, and I do not want to take the risk of it being used for other purposes.*" (P18)). However, such a concern can be alleviated if the storytellers provide more detailed explanations, as noted by P13, "*at the beginning of the story, I was asked to fill in my age, address, and occupation without any explanation. Now I was reluctant to do so but I think more contextual information would help.* When involving the input of knowledge and wisdom, the participants usually felt less concerned about the risk of information privacy. Rather they often had an enthusiasm to participate in the interaction to "test" their knowledge or wisdom (e.g., "*I was happy when I guessed several questions correctly, it shows I have a keen eye for what is happening around me*" (P11)).

Interactivity-Comprehensibility Balance The distribution of the ratings suggested that for C1-C6, *the balance between interactivity and comprehensibility* is kept well. Specifically, the six patterns significantly improved the interactivity of data stories while did not hinder readers from understanding the stories. Overall, more than half of the participants said that the interactivity empowered by BTFW interaction helps them to learn the storytellers' messages through the lens of their personal experience (e.g., "*Based on the address I entered, the system analyzed the vaccine situation in my location from different dimensions, which gave me a better understanding of the reasons for the current vaccine shortage*" (P11)). Besides, the participants appreciated the ability to freely explore the story given by BTFW interaction, such as details on demand and zoom. In this way, they were able to learn more data insights of interest. For example, P15 said, "*what I really need to learn is the capacity of the hospitals around me rather than in the entire city, so I can have a clear view about the health care resource accessible to me*". As for comprehensibility, although C1-C6 were all considered easier to understand than C0, C6 was relatively the least comprehensible one. Some participants explained that it was because C6 fails to provide adequate explanation beyond the interaction, which left barriers for readers to understand the content (e.g., "*I was hoping to learn the reason for the price change, however, the author ended the story without any explanation after I answered all the questions*" (P2)).

The Learning Curve of Interaction In terms of the challenge concerning the learning curve on interaction mentioned by the experts in our preliminary study, the participants experienced a relatively flat learning curve when interacting with C1-C6. This finding may be related to two reasons: on the one hand, the overall age of the participants is relatively young and they generally have a lot of experience interacting with digital media (e.g., "*The interactions in these cases are used frequently in my daily life, so I was able to understand the interaction rules set by the storyteller*" (P9), "*When I saw the slider bar I knew that it could be used to adjust the value of R0 immediately even before I understood what does R0 mean*" (P16)); On the other hand, we intentionally selected stories with intuitive interaction design as stimuli, such as click and scrolling. For example, P7 commented, "*the interaction techniques applied in these stories are quite fundamental, and I have used interaction techniques with much more complex mechanisms.*" Also, how interaction is designed can affect the steepness of the learning curve, e.g., if its perceived affordance was intuitive, and if it provides timely feedback on user input. For example, P21 said, "*C3 provides clear visual cues of interaction, which can be performed via three steps: it first allows me to adjust the parameters, then it starts calculating the values, and the final step presents the visualization of the results. During the interaction, it was clear to me what I was supposed to do at each step.*" while P19 noted, "*one story only shows the results of my input on the second half of the article. When interacting with it at the beginning, I thought I encountered a bug since it did not respond to my input in real-time.*" However, several participants complained about the difficulty they met in manipulating C5,"*I mistook the bar chart which indicates the percentage of the cause of death for a slider. I tried to slide it but failed many times, I'm really frustrated.*" (P14).

In summary, the qualitative analysis of the participants' responses suggested that BTFW interaction can effectively augment *self-story connection, user engagement*, and *information recall*, which are in line with the findings derived in the expert interviews. However, contrary to what the experts expected, only a few participants mentioned that they had concerns about using BTFW interaction, including *information privacy, interactivity-comprehensibility balance*, and *the learning curve of interaction*. The reason may be that we carefully selected the stimuli for our user study. We will discuss how to address these concerns in the following section.

7 Discussion

In this section, we discuss design implications about BTFW interaction derived from the analysis of our corpus and the observation of our user study. We also discuss the limitations of our current work.

7.1 Possibilities for Designing BTFW Interaction

According to the coding of our corpus and the result of the user study, we observed three possibilities for designing BTFW interaction. The first design implication is the Establisher's superiority over other narrative categories in integrating with BTFW interaction according to the results of our user study. The main reason may be that it is easier to build readers' connections at the beginning of stories. Therefore, we suggest that for storytellers who attempt to design BTFW interaction for their data stories, embedding its output in the Establisher is a good choice to consider.

Second, it is important to encourage various input modalities of BTFW interaction. The existing input modalities found in our corpus mainly use mouse-based interaction, including selecting, typing, sliding, and drawing. The input modality can have a profound impact on the reading experience. For example, after exploring "*You draw it*" [70], one participant (P27) commented that "*drawing is quite a novel interaction technique for data stories, and its high degree of freedom promoted me to get greater enjoyment*". We suggest that more modalities such as speech-based interaction can be integrated into data stories. For example, through natural language interaction, readers can use flexible expressions (e.g., "Days I exercised for more than 10 minutes") to explore data [39]. In addition, input taking the form of speech can reflect more characteristics such as affective prosody (e.g., changes in pitch, loudness, and speech rate) in readers.

Third, it is promising to explore more possibilities focusing on the blank cells in our framework that are untouched by BTFW interaction (Fig. 3). For example, no stories in our corpus are found to use BTFW interaction that requires readers' knowledge and responds to it in the Initial. As we have explained in Section 4, combinations of input and output in different levels can have various impacts on the reading experience. Thus, exploring the blank cells in our framework may lead us to find potential benefits of BTFW interaction such as evoking affective responses or motivating action.

7.2 Challenges for Designing BTFW Interaction

To address the challenges proposed in both the preliminary study and the user study, we also derived three design implications. First, to address privacy concerns, we suggest using fuzzy input instead of exact input. In our corpus, a few interactive data stories employ such a strategy. Using fuzzy input allows readers to get personalized content while blurring their personal information. For example, in "*Should prison sentences be based on crimes that haven't been committed yet?*" [7], readers can choose their age groups (e.g., $<$25, 25-44, and 45+) rather than inputting their exact ages. In "*How many pain pills went to your pharmacy?*" [31], readers can search for the name of nearby pharmacy stations rather than the exact residential address to check the situation around.

Second, we look at the longstanding debate on whether or not to integrate interaction into data stories [71] through the lens of BTFW. Thus, we suggested that adding an appropriate level of interactivity to data stories can improve the reading experience of data stories. For example, simple interaction techniques such as text-based search [32] and sketch-based prediction [40] provide an intuitive method to explore and explain data. Considering the upcoming trend of consuming data stories in immersive devices [55, 65], we believe that interaction will increasingly play a more important role in telling stories with data in the future.

The third suggestion is to ease the learning curve of BTFW interaction by following the widely recognized principle in HCI, affordance [52], which describes the relationship between what an object looks like and how it is used. While designing BTFW interaction, storytellers can use various visual cues to imply that an interaction is possible. For example, adding a mark that attracts attention to the interactive area to suggest it can be clicked [14]. In interactive data stories, managing affordances is critical, users often perceive possible actions based on the properties of interface or visualization elements. Using incorrect visual cues can hinder user experience and lead to frustration.

7.3 Limitations and Future Work

There are several limitations of our work. First, the six design patterns of BTFW interaction are derived from the analysis of our corpus consisting of 58 data stories. As the corpus is not exhaustive nor representative of the field, the design patterns are considered as an initial step toward understanding interaction in data stories. Therefore, we encourage narrative visualization professionals and practitioners to expand the corpus and explore more BTFW interactions by referring to our framework. Second, when selecting stimuli for our user study, we used one specific example data story from each design pattern as its representative. Thus, it is difficult to draw a generalizable conclusion from the Likert ratings of individual point examples from our corpus. Also, our study stimuli are constrained to the topics of Covid-19. Covid-19 is a global event that concerns everyone's life. This is the reason why most of the participants can connect to this topic even without the promotion from BTFW interaction. A follow-up study could examine topics that are not directly relevant to readers, such as the origin of the universe, to gain a deeper understanding of what benefits can BTFW interaction bring. Also, in five out of seven stimuli, their data are mainly about the United States. The measurement of self-story connection and engagement can be affected if the data itself does not have any connection with the participants who are not from the US. Thus, to further understand the use of BTFW interaction, future studies can use stimuli that focus on less culture-related topics. Third, the participants we recruited are all relatively well-educated. Considering the relationship between education level and visualization literacy [54], further work that focuses on user groups with different educational backgrounds can be conducted.

Our future work plans to explore more possibilities of combining BTFW with data-driven storytelling. Our current work focuses on BTFW that directly addresses the audience. However, BTFW can also be achieved by having characters in a fictional story recognize their own fictitious existence. A common example is pulling back the camera in a sitcom and allowing the audience to see the walls that contain the scene. In this sense, the equivalent of data-driven storytelling would be a data story explaining the construction of specific visualizations (e.g., *What's so hard about histograms?* [47]), which are usually hidden from readers. Another example that calls attention to the artificiality of the medium is when an actor is corpsing, that is, he or she breaks character during a scene by laughing or forgetting their lines. Similarly, such corpsing happens to data stories if glitches in visualizations are exposed to readers, due to bugs or bit rot (e.g., *The bits are rotting in the state of data journalism* [42]).

8 Conclusion

This paper introduces the concept of breaking the fourth wall (BTFW) in the context of data-driven storytelling and explores different design patterns of BTFW interaction. The results of our user study suggested that BTFW interaction has the potential to enhance self-story connection, user engagement, and information recall. On the other hand, BTFW interaction should be carefully used considering the balance between interactivity and comprehensibility, information privacy, and the learning curve of interaction. Also, we proposed design implications to address the possibilities as well as the challenges for designing BTFW interaction. We consider this work as a first step toward understanding interaction in data-driven storytelling through the lens of breaking the fourth wall. We hope our findings can promote the integration between data stories and storytelling devices for designing expressive data stories.

Acknowledgments

Nan Cao is the corresponding author. This work was supported by NSFC 62072338 and NSF Shanghai 20ZR1461500. We would like to thank all the reviewers for their constructive feedback.

REFERENCES

[1] R. L. Ackoff. From data to wisdom. *Journal of applied systems analysis*, 16(1):3–9, 1989.

[2] 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 2015 CHI Conference on Human Factors in Computing Systems*, pp. 1459–1468. ACM, 2015.

[3] B. Anderson and E. Borges-Rey. Encoding the ux: User interface as a site of encounter between data journalists and their constructed audiences. *Digital Journalism*, 7(9):1253–1269, 2019.

[4] P. J. Auter. Psychometric: Tv that talks back: An experimental validation of a parasocial interaction scale. *Journal of Broadcasting & Electronic Media*, 36(2):173–181, 1992.

[5] P. J. Auter and D. M. Davis. When characters speak directly to viewers: Breaking the fourth wall in television. *Journalism Quarterly*, 68(1-2):165–171, 1991.

[6] B. Bach, D. Stefaner, J. Boy, S. Drucker, L. Bartram, J. Wood, P. Ciuccarelli, Y. Engelhardt, U. Koeppen, and B. Tversky. Narrative design patterns for data-driven storytelling. In *Data-Driven Storytelling*, pp. 107–133. AK Peters/CRC Press, Boca Raton, FL, US, 2018.

[7] A. M. Barry-Jester, B. Casselman, and D. Goldstein. Should prison sentences be based on crimes that haven't been committed yet? `https://fivethirtyeight.com/features/prison-reform-risk-assessment/`, 2015. Accessed: 2022.

[8] BBC. Bmi and obesity: Where are you on the uk fat scale? `https://www.bbc.com/news/health-43697948`, 2018. Accessed: 2022.

[9] E. Bell. *Theories of performance*. Sage, 2008.

[10] A. Bhatia and J. Katz. Why we're experiencing so many unusually hot summer nights. `https://www.nytimes.com/interactive/2021/09/13/upshot/up-hot-nights.html`, 2021. Accessed: 2022.

[11] Bloomberg. Find out if your job will be automated. `https://www.bloomberg.com/graphics/2017-job-risk/`, 2017. Accessed: 2022.

[12] I. Bogost, S. Ferrari, and B. Schweizer. *Newsgames: Journalism at play*. Mit Press, 2012.

[13] L. Bonet and E. Négrier. Breaking the fourth wall: Proactive audiences in the performing arts. *Space*, 26(7.1):44–9, 2018.

[14] J. Boy, L. Eveillard, F. Detienne, and J.-D. Fekete. Suggested interactivity: Seeking perceived affordances for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):639–648, 2015.

[15] V. Braun and V. Clarke. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2):77–101, 2006.

[16] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE transactions on visualization and computer graphics*, 19(12):2376–2385, 2013.

[17] O. G. Brockett. *The theatre: An introduction*. Holt, Rinehart and Winston, 1974.

[18] C. Brontë. Jane eyre. In *Medicine and Literature*, pp. 53–72. CRC Press, 2018.

[19] T. Brown. *Breaking the fourth wall*. Edinburgh University Press, 2013.

[20] B. Brummett. *Techniques of close reading*. Sage Publications, 2018.

[21] D. Cartmell and I. Whelehan. *The Cambridge companion to literature on screen*. Cambridge University Press, 2007.

[22] M. Cavazza and D. Pizzi. Narratology for interactive storytelling: A critical introduction. In *International Conference on Technologies for Interactive Digital Storytelling and Entertainment*, pp. 72–83. Springer, 2006.

[23] N. Cohn. Visual narrative structure. *Cognitive science*, 37(3):413–452, 2013.

[24] M. Conlen, J. Keefe, A. Sun, L. Leatherby, and C. Smart. How full are hospital i.c.u.s near you? `https://www.nytimes.com/interactive/2020/us/covid-hospitals-near-you.html`, 2020. Accessed: 2022.

[25] S. Conway. A circular wall? reformulating the fourth wall for videogames. *Journal of Gaming & Virtual Worlds*, 2(2):145–155, 2010.

[26] J. A. Cuddon. *A dictionary of literary terms and literary theory*. John Wiley & Sons, 2012.

[27] P. Dawson, R. Levy, and N. Lyons. 'breaking the fourth wall': 3d virtual worlds as tools for knowledge repatriation in archaeology. *Journal of Social Archaeology*, 11(3):387–402, 2011.

[28] N. DiCamillo. Guess how much prices on 11 common grocery items have risen in one year. `https://qz.com/2088955/us-food-prices-are-rising-take-a-quiz-to-find-out-by-how-much/`, 2021. Accessed: 2022.

[29] E. Dimara and C. Perin. What is interaction for data visualization? *IEEE transactions on visualization and computer graphics*, 26(1):119–129, 2019.

[30] B. Dykes. *Effective data storytelling: how to drive change with data, narrative and visuals*. John Wiley & Sons, 2019.

[31] A. Emamdjomeh, A. Ba Tran, A. Williams, D. Rindler, Meko, and Tim. How many pain pills went to your pharmacy? `https://www.washingtonpost.com/graphics/2019/investigations/pharmacies-pain-pill-map/`, 2020. Accessed: 2022.

[32] M. Feng, C. Deng, E. M. Peck, and L. Harrison. The effects of adding search functionality to interactive visualizations on the web. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2018.

[33] G. Freytag and E. J. MacEwan. *Freytag's technique of the drama: an exposition of dramatic composition and art*. Scott, Foresman and Company, 1908.

[34] J. Heer and B. Shneiderman. Interactive dynamics for visual analysis. *Communications of the ACM*, 55(4):45–54, 2012.

[35] J. Heyer, N. K. Raveendranath, and K. Reda. Pushing the (visual) narrative: the effects of prior knowledge elicitation in provocative topics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.

[36] F. Hohman, M. Conlen, J. Heer, and D. H. P. Chau. Communicating with interactive articles. *Distill*, 5(9):e28, 2020.

[37] K. Hornbæk and A. Oulasvirta. What is interaction? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 5040–5052, 2017.

[38] J. Interlandi and Y. Serkez. What are the vaccine roadblocks where you live? `https://www.nytimes.com/interactive/2021/02/25/opinion/covid-vaccination-barriers.html`, 2021. Accessed: 2022.

[39] Y.-H. Kim, B. Lee, A. Srinivasan, and E. K. Choe. Data@ hand: Fostering visual exploration of personal data on smartphones leveraging speech and touch interaction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021.

[40] Y.-S. Kim, K. Reinecke, and J. Hullman. Explaining the gap: Visualizing one's predictions improves recall and comprehension of data. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1375–1386, 2017.

[41] M. Koerth and E. Mejía. Why even a small thanksgiving is dangerous. `https://fivethirtyeight.com/features/why-even-a-small-thanksgiving-is-dangerous/`, 2020. Accessed: 2022.

[42] R. Kosara. The bits are rotting in the state of data journalism. `https://eagereyes.org/blog/2016/the-bits-are-rotting-in-the-state-of-data-journalism`, 2016. Accessed: 2022.

[43] X. Lan, Y. Wu, Y. Shi, Q. Chen, and N. Cao. Negative emotions, positive outcomes? exploring the communication of negativity in serious data stories. In *In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2022.

[44] X. Lan, X. Xu, and N. Cao. Understanding narrative linearity for telling expressive time-oriented stories. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.

[45] B. Lee, N. H. Riche, P. Isenberg, and S. Carpendale. More than telling a story: Transforming data into visually shared stories. *IEEE Computer Graphics and Applications*, 35(5):84–90, 2015.

[46] D. Lu. Here's every total solar eclipse happening in your lifetime. is this year your best chance? `https://www.washingtonpost.com/graphics/national/eclipse/`, 2017. Accessed: 2022.

[47] A. Lunzer and A. McNamara. What's so hard about histograms? `https://tinlizzie.org/histograms/`, 2020. Accessed: 2022.

[48] Lupa. What if all covid-19 deaths in brazil happened in your neighborhood? `https://piaui.folha.uol.com.br/lupa/epicentro/en/`, 2020. Accessed: 2022.

[49] S. McKenna, N. Henry Riche, B. Lee, J. Boy, and M. Meyer. Visual narrative flow: Exploring factors shaping data visualization story reading experiences. *Computer Graphics Forum*, 36(3):377–387, 2017.

[50] J. Meyrowitz. Television and interpersonal behavior: Codes of perception and response. *Inter/media: Interpersonal Communication in a Media World*, 253:272, 1986.

[51] J.-E. Nordlund. Media interaction. *Communication Research*, 5(2):150–175, 1978.

[52] D. A. Norman. Affordance, conventions, and design. *interactions*, 6(3):38–43, 1999.

[53] Osserman. People of the pandemic. https://peopleofthepandemicgame.com/, 2020. Accessed:2022.

[54] E. M. Peck, S. E. Ayuso, and O. El-Etr. Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2019.

[55] D. Ren, B. Lee, and T. Höllerer. Xrcreator: interactive construction of immersive data-driven stories. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, pp. 1–2, 2018.

[56] E. F. Risko, D. C. Richardson, and A. Kingstone. Breaking the fourth wall of cognitive science: Real-world social attention and the dual function of gaze. *Current Directions in Psychological Science*, 25(1):70–74, 2016.

[57] K. Rogers, J. Wolfe, and L. Bronner. Without a vaccine, herd immunity won't save us. https://fivethirtyeight.com/features/without-a-vaccine-herd-immunity-wont-save-us/, 2020. Accessed: 2022.

[58] H. Romat, N. Henry Riche, C. Hurter, S. Drucker, F. Amini, and K. Hinckley. Dear pictograph: Investigating the role of personalization and immersion for consuming and enjoying visualizations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2020.

[59] R. E. Roth. An empirically-derived taxonomy of interaction primitives for interactive cartography and geovisualization. *IEEE transactions on visualization and computer graphics*, 19(12):2356–2365, 2013.

[60] J. Rowley. The wisdom hierarchy: representations of the dikw hierarchy. *Journal of information science*, 33(2):163–180, 2007.

[61] H.-J. Schulz, T. Nocke, M. Heitzler, and H. Schumann. A design space of visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2366–2375, 2013.

[62] K. Sedig and P. Parsons. Interaction design for complex cognitive activities with visual representations: A pattern-based approach. *AIS Transactions on Human-Computer Interaction*, 5(2):84–133, 2013.

[63] E. Segel and J. Heer. Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1139–1148, 2010.

[64] 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 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2021.

[65] D. Shin and F. Biocca. Exploring immersive experience in journalism. *New Media & Society*, 20(8):2800–2823, 2018.

[66] C. D. Stolper, B. Lee, N. Henry Riche, and J. Stasko. Emerging and recurring data-driven storytelling techniques: Analysis of a curated collection of recent stories. Technical Report MSR-TR-2016-14, Microsoft, April 2016.

[67] N. Sultanum, F. Chevalier, Z. Bylinskii, and Z. Liu. Leveraging text-chart links to support authoring of data-driven articles with vizflow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021.

[68] P. R. Teske and T. Horstman. Transmedia in the classroom: Breaking the fourth wall. In *Proceeding of the 16th International Academic MindTrek Conference*, pp. 5–9, 2012.

[69] S. Thompson and J. Colombo. Find your place in the vaccine line. https://www.nytimes.com/interactive/2020/12/03/opinion/covid-19-vaccine-timeline.html?action=click&module=Opinion&pgtype=Homepage, 2020. Accessed: 2022.

[70] T. N. Y. Times. You draw it. https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html, 2015. Accessed: 2022.

[71] A. Tse. Why we are doing fewer interactives. *Malofiej Infographics World Summit*, 2016.

[72] M. Wallis and S. Shepherd. *Studying plays*. Bloomsbury Publishing, 2018.

[73] Z. Wang, H. Romat, F. Chevalier, N. H. Riche, D. Murray-Rust, and B. Bach. Interactive data comics-vis2021. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[74] P. Waugh. *Metafiction: the theory and practice of self-conscious fiction*. Routledge, 2002.

[75] E. Wilson and A. Goldfarb. *Living theatre: history of the theatre*. McGraw-Hill Boston, MA, 2008.

[76] N. Yau. How you will die. https://flowingdata.com/2016/01/19/how-you-will-die/, 2016. Accessed: 2022.

[77] J. S. Yi, Y. ah Kang, J. Stasko, and J. A. Jacko. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1224–1231, 2007.

[78] Q. Zhi, A. Ottley, and R. Metoyer. Linking and layout: Exploring the integration of text and visualization in storytelling. In *Computer Graphics Forum*, vol. 38, pp. 675–685. Wiley Online Library, 2019.