Coupling Interaction: Engaging Users in Narrative Visualization by Linking Text and Visualization

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

Leave Authors Anonymous
for Submission
City, Country
e-mail address

ABSTRACT

While narrative visualizations are popular in news media outlets, the interplay between text and visuals in an interactive context remains underexplored. We define coupling interaction, a bidirectional interaction mode that explicitly links narrative text with its explanatory visualization element and vice versa. We explore the effects of integrating coupling interaction and varying layout settings on the narrative visualization reading experience. Through a crowdsourced study with 250 participants, we find that coupling interaction can significantly increase user-engagement for narrative visualizations. While participant feedback indicates that they prefer the slideshow layout and coupling interaction, we did not find evidence that either improves story comprehension or recall performance. We also visualize and analyze interaction behavior data that was recorded during participants' reading of the narrative visualization. Based on our findings, we suggest design strategies for how to integrate coupling interaction into narrative visualization.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI)

Author Keywords

Narrative visualization; storytelling; coupling interaction; crowdsourcing; engagement; evaluation

INTRODUCTION

Narrative visualizations that integrate visual representations into narratives have been used to reveal and convey data-rich stories [32]. For example, news media incorporate graphics to provide supporting evidence or related details in a compelling format. Analysts embed visualizations into their reports to communicate their findings in support of decision making [21]. This has led the visualization community to develop analytic frameworks to inform design strategies for narrative visualization [18, 32], study how narrative flows [28] and storytelling [6] affect the story reading experience, and implement tools [19, 31] to support the creation of narrative visualizations. Nonetheless, a deeper understanding of the effect of interaction and layout on the narrative visualization reading experience remains elusive.

The essence of information visualization (InfoVis) is interactivity and exploration [6]. A single visualization typically provides answers to, at best, a handful of questions. Instead, visual analysis with interaction allow user to uncover insights with an iterative process of view creation, exploration, and refinement [16]. However, interaction rarely is the main focus of InfoVis research [36]. In narrative visualization, the story appear to be more effective with constrained interaction as interactivity might detract from the author's intended message [32]. Thus, finding ways to integrate interaction into narrative visualization and allow users to gain insight by exploration can prove beneficial.

Here, we explore the potential of coupling interaction to trigger the desired user engagement and help users understand the narrative visualization. Coupling interaction is defined as a bidirectional interaction mode that explicitly links narrative text with its explanatory visualization element and visualization with its corresponding narrative text. We conduct a crowdsourcing study with 250 participants to assess the effect of coupling interaction on story comprehension, recall, and engagement. We also investigate how different narrative visualization layout settings impact the story reading experience. We study which narrative visualization layout is most appropriate for adding coupling interaction. We further visualize and analyze the user interaction behavior data and suggest design strategies for using coupling interaction. As a result, we:

- demonstrate that coupling interaction improves the overall user-engagement.
- find participants had a subjective preference for the slideshow layout, but there is no evidence that any of the three has a significant advantage with regards to objective metrics including story comprehension, recall, and engagement.
- demonstrate adding coupling interaction is more helpful in narrative visualization with vertical layout and the side-by-side layout.
- summarize the different user interaction behavior for three narrative visualization layouts (vertical, side-by-side, slideshow) and investigate the impact of coupling interaction on interaction behavior.

CHI'18, April 21–26, 2018, Montreal, CANADA

DOI: 10.1145/1235

BACKGROUND

In this section, we discuss previous research on narrative visualization, the use of interaction and layout in narrative visualization, evaluation methods, and reading behavior with respect to narrative visualization.

Narrative Visualization

Technology provides us with new media and genres that can now be used to convey information in a story-like fashion [13]. Narrative visualization that integrate storytelling fashion into information visualization is a popular research area in visualization community and promises to open up entirely new avenues of research in visualization [21]. Segal et al. proposed a design space for narrative visualization, identified seven *genres*, and proposed three structures for balancing author-driven and reader-driven stories: Martini Glass, Interactive Slideshow, and Drill-Down Story [32]. Hullman et al. also proposed an analytical framework named *"visualization rhetoric"* for understanding how design techniques that prioritize particular interpretations in visualizations that *"tell a story"* can significantly affect end-user interpretation.

These frameworks have been widely used for narrative visualization design. Graph comics were used as a medium to communicate changes in dynamic networks [3]. Boy et al. studied using narratives for engaging users to explore data [6]. Narrative visualizations were also designed for online journalists [32] and science communication [24]. Many tools have been proposed to support narrative visualization design [23, 31, 19]. Studies [20, 28] also have been conducted to understand how different design factors affect the reading experience of visual stories. Building on these findings, we further study how adding interaction and different layout settings impact narrative visualization story reading experience.

Interaction in Narrative Visualization

Interaction is one of the most important parts of visualization: being able to not just see the data, but quickly change the view, add different data, etc., makes analyzing it much faster and more effective [21]. Successful interactive visualization combines expressive graphical representations and effective user interaction [34], yet the information visualization community has generally focused more on visual encoding rather than on interaction. To understand the cost of using interaction, Lam et al, proposed a framework that includes seven costs for interaction to help visualization designers weigh the costs and benefits of interaction based on empirical results [22]. Heer and shneiderman also presented a taxonomy of interactive dynamics that contribute to successful analytic dialogues [16].

In narrative visualization, data stories appear to be most effective when they have constrained interaction because interactivity might engage the user but detract from the author's intended message as well [32]. It is still unclear how to incorporate interaction into a story flow without becoming a distraction, and without the interaction interfering with the flow and point of the story [21].

Layout in Narrative Visualization

Layout in narrative visualization is the way in which text and graphs are set out in the story. Martinec et al. presented a generalized system of image-text relations which applies to different genres of multimodal discourse in which images and texts co-occur [27]. For example, they proposed the relation between text and pictures can be independent or complementary. Segal et al. characterized narrative visualizations into seven basic *genres*: magazine style, annotated chart, partitioned poster, flow chart, comic strip, slide show, and film/video/animation. These *genres* vary in terms of ordering of visual elements, which also implies the layout settings. However, it remains underexplored that how different layout settings impact the narrative visualization story reading experience. In this paper, we adapted a narrative visualization into three layout settings: vertical, side-by-side, slideshow, and investigate the effects on story reading.

Evaluating Reading Experiences

Telling a story via narrative visualization appears to be effective for conveying the intended message. However, there are no clearly defined metrics or evaluation methods to measure the effectiveness [21]. To investigate the effect of adding embellishments into visualization, Bateman et al. measured participants' interpretation accuracy and long-term recall by asking some description questions [4]. Dimara et al. explored the effects of providing task context when evaluating visualization tools using crowdsourcing [9]. They evaluate the comprehension and recall by measuring participants' ability to perform and understand some tasks. To evaluate the comprehension and recall of a narrative visualization, we adapted these methods into our study and design some specific tasks to measure participants' understanding.

Engagement also plays an important role in determining the effectiveness and impact of narrative visualizations [28]. Engagement is a complex topic because it lacks a unified definition in the community [25]. Boy et al. defined engagement as user's investment in the exploration of a visualization [6]. They evaluated engagement by analyzing user-semantic operations and depth of interaction. Time spent on subjective reaction cards for capturing user feelings were also measured to assess user experience [30]. O'Brien et al. defined and validated a subjective questionnaire for measuring engagement based on attributes such as focus attention, perceived usability, aesthetics, endurability, novelty, and felt involvement [29]. In this paper, we evaluate engagement with the combination of reader's subjective feedback and the investment in the exploration of a narrative visualization. We adapted the 14 subjective questions, which originated from O'Brien et al. [29] and tailored by McKenna et al.[28]. Additionally, we recorded the depth of interaction, which is interpreted as the number of interaction a user performs [6], as an objective metric for engagement evaluation.

User Reading Behavior

A promising research direction of narrative visualization is to focus on readers' experiences when viewing and interacting with narrative visualizations [32]. Eye-tracking techniques were employed to study newspaper reading patterns [12, 17]. They found readers regularly skim by scanning graphics, headlines, and initial paragraphs before read an article. Garcia et al. noted that when designing a newspaper, we need to

Figure 1. The study interfaces for coupling interaction study conditions with three different layout settings: vertical, side-by-side, slideshow. It shows that the explanatory visualization elements are highlighted when participants hovering over the first sentence in section "Some Have Prospered".

find readers material that is worthy of their scan, that makes them stop scanning and start reading [12]. In this paper, we recorded interaction behavior data during participants' reading of the narrative visualization. Based the collected data, we compared the viewing behavior for different layout settings and analyzed if adding coupling interaction can change the viewing and interacting behavior.

RESEARCH QUESTIONS

We framed four research questions to guide the study design. Specifically, our research questions were:

- **RQ1: Does adding coupling interactions help?** Coupling interaction is defined to help users comprehend and explore narrative visualizations. The purpose of this question was to examine the effect of coupling interaction on the overall reading experience in terms of engagement, comprehension, recall, and subjective metrics.
- **RQ2: Which layout is most appropriate for narrative visualizations that comprise of individual sections?** The narrative visualization used in this paper contains several independent sections where each section includes narratives and visualizations. The purpose of this question was to understand which layout setting is most appropriate for this narrative visualization category.
- **RQ3: Of the three layout settings, which is the most appropriate for coupling interaction?** The purpose of this question was to compare the three different layout settings and find which one has the largest effect on adding coupling interaction.
- **RQ4: How does the user interaction behavior differ for different layout settings or adding coupling interaction?** This question was for investigating users' interaction behavior on the narrative visualization. The behavior might differ for different layout settings. Adding coupling interaction might also impact the interaction behavior.

CROWDSOURCING STUDY

In this section, we first defined coupling interaction. Then we described a large scale between-subjects crowdsourced study with 250 participants using Amazon Mechanical Turk (AMT) to understand how coupling interaction and different layout settings affect the narrative visualization reading experience.

Coupling Interaction

Interaction allows users to effectively uncover the information visualization insights. However, a potential problem with interaction is that it might "detract from the author's intended message" [32]. Consider a reader who seeks the explanatory visualization as she reads the narrative text. Traditionally, she would need to comprehend the visualization first, then do the filter or other interactions and go back to the narratives. Instead, we ask: how can we decrease the distraction of the switching process and help readers comprehend and engage with the narrative visualization? In existing narrative visualizations, we found interactions including zooming, filtering, and searching are limited to the scope of the visualizations. The somewhat recent Washington Post interactive visualization *"A visual guide to 75 years of major refugee crises around the world"* [8] is a rare exception that links narrative with its corresponding visualization element by enabling a user to hover over text and explore the graph, however, it fails to link visualization back to the narrative. To explore this opportunity, we defined coupling interaction, a novel interaction mode that links narrative with its explanatory visualization element and visualization with its corresponding narrative in visual stories. In the rest of this section, we explored whether adding coupling interaction can improve story comprehension, recall, and engagement through a large scale crowdsourcing study.

Participants

In total, we recruited 250 participants with a maximum $2.20 reward from AMT. Each participants had at least a 98% HIT approval rate, at least 100 approved HITs, was at least 18 years old, and was from the United States. We excluded two participants for participating multiple times, and one participant who claimed visualizations in the story "messed up". Participants had a varied education background: 15% had a master's or advanced degree, 43.3% had a bachelor's degree, 41.7% had high school or some college experience, 44.9% were female, 54.7% were male (one participant claimed the sex as "other"), nine participants were immigrants, and only 4% were previously familiar with the subject matter presented in the story.

Study Condition Settings

We utilized the narrative visualization *"Immigrants From Banned Nations: Educated, Mostly Citizens and Found in Every State"* by Ford Fessenden et al. of the New York Times [11]. We selected this story because it had many separate sections with each section including both narrative and visualizations, which makes it suitable for many layouts including the slideshow layout. This story also features political insights and associated visualizations, thus we believed that readers would be encouraged to explore both the narratives and visualizations of the story.

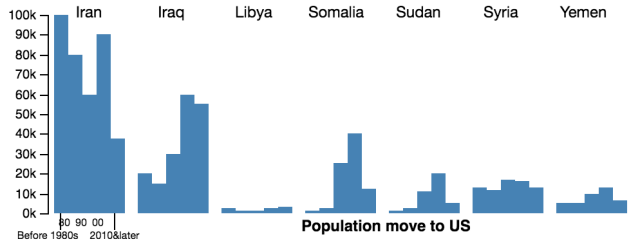

Figure 2. The "population move to US" chart in the narrative visualization. It shows the arrival pattern of immigrants from the seven countries.

task	question
Vis Ext	Of the countries mentioned in the article, which country has the highest immigration population in the US?
TextVis Com	Of the two countries mentioned in the article where the immigrants are better educated than the rest of America, which year accounts for the largest population move to US?

Table 1. Two example questions. The first question is a Vis and Ext task, the second is a TextVis and Com task.

Segal et al. [32] characterized seven different basic *genres* of narrative visualization. To study how different layout settings affect the reading experience, we adapted the story into three different layout settings: vertical (magazine style), side-by-side (similar to the partitioned poster) and slideshow as shown in Figure 1.

Additionally, to better understand the effects of coupling interaction, we created coupling interaction condition and without interaction condition for each layout setting. In conditions with coupling interaction, when a participant hovered over an underlined sentence, as Figure **??** shows, the explanatory visualization elements related to this sentence was highlighted. The corresponding underlined sentence was also highlighted if the participant hovered over an element, such as a bar or a map, of a visualization that corresponded to that particular text element. We configured mouse sensitivity and time interval with HoverIntent [1] to avoid the accidental hover over event.

In total, we have three layout settings with each layout setting has two variations resulting in six conditions:

- **Original** (Original) condition (baseline), where visualizations are individually embedded into narratives with a vertical layout setting.
- **Original with interaction** (OriginalInter) condition (as shown in the first figure in Figure 1), where the narrative visualization is presented by a vertical layout with coupling interaction.
- **Side-by-side** (Side) condition, where we put all narratives on the left and visualizations on the right.
- **Side-by-side with interaction** (SideInter) condition (as shown in the second figure in Figure 1), where we added coupling interaction into the narrative visualization with side-by-side condition. It should be noted that in this condition, since visualizations are in loose order, we added an auto-scrolling feature that automatically navigated readers to the corresponding narrative text when they hover over a visualization element.
- **Slideshow** (Slide) condition, where each section in the narrative visualization is presented as an individual slide with narratives on the left and visualizations on the right. The buttons at the bottom can be clicked to navigate.
- **Slideshow with interaction** (SlideshowInter) condition (as shown in the third figure in Figure 1), where we integrated coupling interaction into the story with slideshow condition.

[1] https://briancherne.github.io/jquery-hoverIntent/

Measures

To assess the effect of coupling interaction and layout setting on reading experience, we used objective performance metrics that measured user-engagement with the narrative visualization (engagement), participants' ability to understand the story (comprehension) and recall the story content (recall), as well as subjective metrics based on self-reported impressions and feedback. In this section, we describe how we designed the tasks and measured performance for comprehension, recall, engagement, and subjective metrics. Complete study materials can be found in the Supplemental Materials [2].

Comprehension

Story comprehension can be viewed as finding a temporally-ordered sequence of situations [14]. Given the structure of narrative visualizations, in each *sequence of situation*, participants needed to fully comprehend not only text, but visualizations as well. So our designed task taxonomies include:

- **Text** task (**Text**), which can be solved solely by reading text in the narrative visualization.
- **Visualization** task (**Vis**), which can be answered solely by observing visualizations in the narrative visualization.
- **Text + Visualization** task (**TextVis**), which can be solved only by combining an understanding of both the text and visualizations.

Additionally, we also adapted two taxonomies of low-level information retrieval tasks [2] for visualization tasks:

- **Extremum** task (**Ext**), where participants had to find the extreme value from visualizations (e.g. find the highest bar in a bar chart).
- **Comparison** task (**Com**), where participants had to compare visualization element across the chart (e.g. compare two bars in a bar chart).

For instance, Figure 2 is a visualization from the narrative visualization we used that shows the trend of immigrants from the seven countries moving into the US. For the first question in Table 1, it can be answered by observing the visualization alone, which means it is a **Vis** task. The question also required participants to find an extreme value, so it is also an **Ext** task. The second question cannot be solved by only observing the visualization. Participants need to find the answer to *"which two counties mentioned in the article where the immigrants are better educated than the rest of America, which year accounts*

[2] Supplemental Materials can be access via: [anonymous for review]

for the largest population move to the US" from the text of the narrative visualization. Thus this question is a **TextVis** task. In addition, it also required participants to compare the height of bars, which indicates it is a **Com** task.

As a result, we had five comprehension tasks that included one **Text** task, three **Vis** tasks, and one **TextVis** task. In **Vis** tasks, we had one **Ext** task and two **Com** tasks. We assigned a binary score of 1 for the correct answer and 0 for all other answers.

Recall

Story recall can be viewed as finding a trajectory consistent with episodic memory constraints [14]. Similar to the comprehension task design, we considered both text and visualizations. Specifically, our designed recall task can be summarized as:

- **Overview** task (**Overview**), where participants had to remember the overall idea and select an aspect that has not been discussed in the story.
- **Detail** task (**Detail**), where participants had to remember some details such as which country has not been mentioned in the story.
- **Visualization** task (**Vis**), where participants had to identify which plot was presented to them in the story.

In total, we used three recall tasks to measure the participants' ability to remember the narrative visualization content. We also assigned a binary score of 1 for the correct answer and 0 for all other answers.

Engagement

Measuring a user's level of engagement in the exploration of media data is a complex matter [6]. As such, we need to find appropriate behavioral proxies that are able to describe an analytical or exploratory intent. Here, our methods to measure engagement were twofold. First, we used low-level user-activity traces as *signals*, and we focus on the *depth of interaction* [6], which means the number of interactions a user performs on the narrative visualization. In this paper, we collected the numbers of *hover over* interaction performed on the narrative visualization for every condition.

Second, O'Brien et al [29] provide a validated questionnaire on six factors: Perceived Usability, Aesthetics, Novelty, Felt Involvement, Focused Attention, and Endurability. We adapted the 14 subjective questions, which originated from O'Brien et al. [29] and tailored by McKenna et al.[28] with each question received a 5-point Likert response from every participant.

Subjective Metrics

We used subjective metrics as a complement to the objective metrics. Specifically, subjective metrics include:

- **Confidence**: After each comprehension and recall task, we asked participants how confident they were in their response. Responses were reported on a 5-point Likert scale.
- **Easiness**: After each comprehension and recall task, we asked participants how difficult the question was. Responses were reported on a 5-point Likert scale.
- **Feedback**: We provided an optional question to collect participants' feedback of the study. The goal was to gather any additional information pertaining to the study, layout setting, and coupling interaction.

Procedure

The study followed a between-subject design. Each participant first received the narrative visualization with a random condition. Due to the random assignment process, sample size per condition ranged from 35 to 58 (planned sample size is 42). After reading an introduction of the study, participants were given a tutorial as a guided tour through the study interface with tooltips. For participants in the coupling interaction condition, we showed them the effect of hovering over narratives and visualizations. For participants in conditions with the slideshow layout setting, we showed them how to navigate the narrative visualization by clicking the buttons. For participants in conditions with the side-by-side layout setting, we pointed out the narratives on the left and how to scroll them, and we pointed out all visualizations on the right.

Participants were asked to carefully read and explore the story. After fully comprehending all narratives and visualizations, they started the questionnaire by clicking a button at the bottom of the story. They first performed five comprehension tasks, and for each task they rated their confidence and task easiness. Then they filled out all demographic information and completed the 14 engagement statements. Next, they were given three recall tasks with confidence and easiness ratings following each task. Finally, participants were given the opportunity to provide general feedback and submit the answers. It should be noted that when answering comprehension tasks, participants were allowed go back to check the story content by clicking a "Go Back to Story" button at the top-right corner of the interface (as shown in Figure 1). However, the button is disabled when they were answering the recall questions. After submission, participants were given a random id to paste into AMT to complete their HIT.

Hypotheses

Coupling interaction enables participants to explore the narrative text with its explanatory visualization element and visualization with its corresponding narratives, which, we suspected, might mitigate the "detract from the author's intended message" [32] problem. Thus we expect that:

H1: Participants in the conditions with coupling interaction will comprehend and recall the story more accurately and engage with story more deeply than participants in conditions without coupling interaction for all of the three layout settings.

Our second research question concerns the effect of different layouts. Our study conditions include three different layout settings: vertical, side-by-side, and slideshow. It should be noted that our layout settings can be mapped to the similar *genres* [32]: vertical as magazine style, side-by-side as the partitioned poster, and slideshow as a slide show. Choosing the appropriate layout depends on a variety of factors, such as the audience, medium, and data. Particular layout settings are clearly preferable in some scenarios. In the narrative visualization we used in the study, the interactive slideshow structure may promote a dialogue between author-driven and reader-driven approaches [32], so we expect that:

H2: Participants in the conditions with the slideshow layout will perform better in terms of comprehension, recall, and subjective metrics than participants in conditions with other layout settings.

For the third research question, we considered in the slideshow and vertical layout, participants can easily link the narrative text and the corresponding visualizations. The side-by-side layout, however, suggests a loose order to its visualizations and mapping to the text [32]. Thus we expect:

H3: Side-by-side layout is most appropriate for adding coupling interaction.

For the final research question, we believed different layout settings and coupling interaction can change the user interaction behavior in terms of the order of the interaction and how they switch views. We expect:

H4: Users will perform a different interaction behavior for different layout settings and adding coupling interaction.

DATA ANALYSIS

Although researchers have proposed several quality control approaches to reject poor crowdsourced jobs [1], we accepted and analyzed all 247 jobs. Different conditions are expected to yield different levels of comprehension, recall, engagement, and subjective metrics, therefore excluding low-quality jobs would bias our results [9]. For each participant, we recorded timestamp, condition name, demographic information (age, sex, education level, if they are immigrants, and if they are familiar with the story topic), five comprehension task responses (each has two 5-point Likert responses for confidence and easiness for the task), 14 5-point Likert responses for engagement statements, three recall task responses (each has two 5-point Likert responses for confidence and easiness for the task), the time spent (including story reading time, questionnaire completion time, and time for completing comprehension tasks, recall tasks, demographic questions, and engagement tasks), each text hover over event (including timestamp and text content), and each visualization element hover over event (including timestamp and visualization element ID).

We performed a linear mixed effects analysis using R [33] and lme4 package [5] to explore H1, H2, and H3. As fixed effects, we either entered coupling interaction or layout settings into the model. As random effects, we had individual differences such as sex and age, time spent, as well as the layout for coupling interaction evaluation and coupling interaction for layout evaluation. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. We report P-values of the model with the effect in question against the model without the effect in question. The R code and detailed analyses can be found in the Supplemental Materials [3].

RESULTS OVERVIEW

Before analyzing our main research questions, we first give an overview of all the results to show all performance details for each condition. We report the sample mean and 95% confidence intervals (CIs) for each condition on comprehension, recall, engagement, time, and subjective metrics in terms of confidence and perceived easiness.

Comprehension

The results of comprehension tasks are shown in Figure 3. Each row represents one condition. The first column shows average scores of all tasks. As we can see in this column, our participants did a fair job (scores 0.5 - 0.6 out of 1). It appears that adding coupling interaction may decrease the accuracy in side and slideshow layout, but the evidence is weak.

Participants did the worst job on **TextVis** task (scores 0.2 -0.4 out of 1). Adding coupling interaction may help participant on **Text** task in original vertical and side-by-side layout and on **TextVis** task in slideshow layout, but the evidence is still weak.

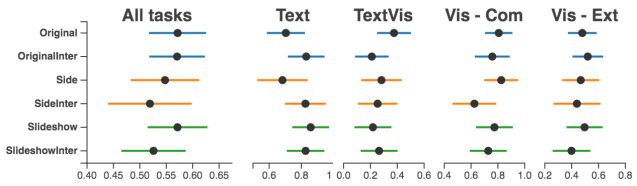

Figure 3. Comprehension tasks result per task taxonomy and condition. Error bars are 95% CIs

Recall

As we can see in Figure 4, although there is no strong evidence, participants in conditions with original vertical layout may outperform others on recall tasks, especially for **Detailed** recall tasks. Adding coupling interaction does not appear to help with recall tasks, which may be due to distraction caused by interaction.

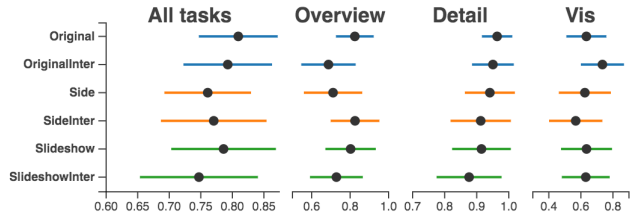

Figure 4. Recall tasks result per task taxonomy and condition. Error bars are 95% CIs

Engagement

In Figure 5, we show the engagement questionnaire scores across six conditions. The result shows increased engagement when reading stories with coupling interaction. There is no strong evidence to show the difference on engagement for different layout settings. A further detailed analysis will be presented in the next section.

We also present the *depth of interaction* at Figure 6. We did not include the interaction data from conditions without coupling

[3] Supplemental Materials can be access via: [anonymous for review]

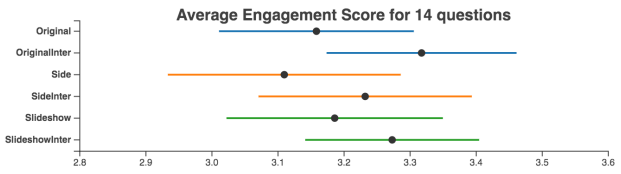

Figure 5. 14 Engagement questions responses for each condition. Error bars are 95% CIs

interaction. It seems participants interacted more with text in side-by-side layout and hovered over more with visualization in vertical layout.

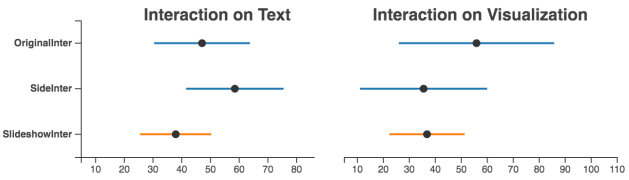

Figure 6. Number of hover over interaction per condition. Error bars are 95% CIs

Subjective Metrics

Figure 7 reports the subjective metrics results. Participants were fairly confident on comprehension and recall tasks (scores 3.5 - 4.0 out of 5). There is no apparent effect on confidence score by adding coupling interaction. The perceived easiness scores are not as high as confidence scores (2.6 - 3.0 out of 5). There is some evidence that participants found tasks in conditions with coupling interaction harder than tasks without interaction.

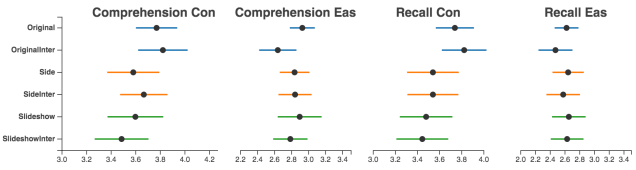

Figure 7. Subjective responses of confidence and easiness on comprehension and recall tasks. Error bars are 95% CIs

Although the general feedback on the study was optional, 81 out of 247 participants provided their opinions. After performing a simple affinity diagram model on the collected feedback, we derived three main themes.

The first theme is related to interaction. Participants in conditions with coupling interaction like the interaction settings. As a participant noted, *"I wish all news articles were interactive like this. It makes it so easy to be able to see the part of the chart that the article is discussing or to see the part of the article that the chart is referencing."* They were also concerned about the possible distraction caused by coupling interaction, as *"I enjoyed the live data highlights that showed up as I read. It was a bit distracting but I think that is just because I am new to this format."*

Participants thought highly of the slideshow layout, as *"I really like the layout of this study. Very easy to navigate and clear directions "* and *"I really liked the reading format/charts. it increased comprehension."*

The final opinion concerned side-by-side layout. Participants tended to express a negative attitude for this layout setting. As some participants said, *"There was WAY too much going on on the story page."* and *"Very difficult story to understand for me"*.

CORE RESULTS

In the previous section, we present an overview of all our results and pointed out several patterns that lack statistical proof. To further explore the data, we used a linear mixed effects model to analyze the results with the purpose of more directly addressing the research questions.

RQ1: Does adding coupling interactions help?
To answer this question, we compared all the performance results between conditions with coupling interaction and conditions without coupling interaction. Specifically, we included all objective and subjective metrics scores in coupling interaction as the fixed effect, with all scores in conditions without coupling interaction as the omitted reference condition. We included layout settings, sex, age, and the time spent as random effects.

We found a significant effect on the engagement performance ($t = 2.05, p = 0.04$), where adding coupling interaction can, on average, increase the engagement score (0.13). However, we observed no evidence that adding coupling interaction has a significant effect on participants' comprehension ($t = -1.03, p = 0.3$), recall ($t = -0.52, p = 0.6$) tasks, and subjective metrics including confidence score ($t = -0.06, p > 0.5$) and easiness score ($t = -1.78, p = 0.08$) on comprehension tasks and confidence score ($t = -0.04, p > 0.5$) and easiness score ($t = -0.88, p > 0.3$) on recall tasks.

In a nutshell, our results partially supported **H1**, and indicated that adding coupling interaction can significantly improve the engagement score. However, there was no measurable benefit for comprehension, recall, and the corresponding subjective metrics.

RQ2: Which layout is most appropriate for visual stories comprised of individual sections?
We performed a linear mixed effects analysis for this question with all objective and subjective metrics scores in different layout settings as the fixed effect, and with interaction, sex, age, and time spent as random effects.

We found no apparent difference of the effect on comprehension ($t = -0.9, p = 0.35$), recall ($t = -1, p = 0.3$), and engagement ($t = 0.07, p > 0.5$) in the three layout settings. However, we found a significant effect on the confidence score of comprehension tasks ($t = -2.59, p < 0.01$) and recall tasks ($t = -3.03, p < 0.005$) for the vertical layout setting. We didn't find significant different impact on easiness score for

comprehension tasks ($t = 0.376, p > 0.5$) or easiness score for recall tasks ($t = 1.022, p > 0.3$).

The results rejected **H2**. Layout doesn't seem to matter for our studied conditions. Also, our findings suggested that participants felt most confident when they read stories in the original vertical layout.

RQ3: Of the three layout styles, which is the most appropriate one to add coupling interaction?
To answer this question, we compared all the objective and subjective metrics for each layout setting, with and without coupling interaction. We split our results into three parts by different layout settings. For each part, we included all metrics scores in coupling interaction as the fixed effect, with all scores in conditions without coupling interaction as the omitted reference condition. We included sex, age, and the time spent as random effects.

For the slideshow layout, we found no significant difference in all metrics by adding coupling interaction. As for side-by-side and vertical layout, we found adding coupling interaction can significantly increase the interaction with text (side-by-side: $t = 3.675, p = 0.016$, vertical: $t = 4.653, p < 0.01$). There was no other significant difference for other metrics.

The results partially supported **H3**, suggesting adding coupling interaction is a benefit for vertical and side-by-side layout by increasing reader's interaction with the narrative text.

RQ4: How does the user interaction behavior differ for different layout settings or adding coupling interaction?

To answer this question, we collected all hover over interaction events performed by participants, including a timestamp and the context information for hovering over text and visualization elements. It should be noted that we also record the interaction data for conditions without coupling interaction for comparison purpose.

Our collected interaction behavior data can be viewed as temporal sequential data, which is often analyzed through visual analytic methods [15, 35, 10]. We present the user interaction behavior data visually as shown in Figure 8.

For an overview of all the conditions without coupling interaction, we observed participants tended to focused on interacting with visualizations in side-by-side layout. Participants switched more frequently in the vertical layout. In the slideshow layout, participants tended to solely focus on either the narrative text or the visualizations.

We then analyzed the effect of adding coupling interaction on interaction behavior. For vertical layout, we observed that adding coupling interaction seems to encourage participants to interact with the text, which agrees with our analysis in **RQ1**. With regards to the side-by-side layout, we noticed an apparent pattern that participants focused on interacting with visualizations in the condition without coupling interaction. Adding coupling interaction greatly changed this pattern; interaction with the narrative text dominates the interaction behavior. As for the slideshow layout, adding coupling interaction may encourage participants to interact more between text and visualizations.

To prove our observation, we performed a linear mixed effects model on interaction with text and visualization. we found participants in conditions with coupling interaction interacted significantly more with narrative text ($t = 5.329, p < 0.01$). However, there was no strong evidence that coupling interaction is able to trigger more interactions with the visualization ($t = -0.079, p > 0.5$). There was also no evidence to show an effect on text interaction ($t = 0.023, p > 0.5$), visualization interaction ($t = -1, p = 0.3$) for different layout settings,

We also performed a linear mixed effects model on user's switch tendency between narrative text and visualizations. As shown in figure 9, adding coupling interaction tends to decrease participants' switch frequency between narrative text and visualization. Nevertheless, we did not find a significant effect on switch frequency for adding coupling interaction ($t = -1.6, p = 0.1$) and changing layout settings ($t = -0.688, p = 0.49$).

The results showed user interaction behavior significantly differ for adding coupling interaction

Our results partially supported **H4**. Participant performed a significant different interaction behavior with text when adding coupling interaction. However, our findings suggested there was no evidence to show a significant effect of layout settings on user interaction behavior.

DISCUSSION

Our results suggest the promises and potential pitfalls of incorporating coupling interaction into narrative visualization.

We first defined coupling interaction as a bi-directional interaction mode that explicitly links narrative text with its explanatory visualization element and visualization with its corresponding narrative text. Current interaction in narrative visualization fails to consider linking the narrative text and explanatory visualizations. Our work is the first to study the effect of coupling interaction on narrative visualization reading experience. We expected coupling interaction would be helpful for comprehending the narrative visualization, improving the recall performance, and engaging readers. In fact, we failed to observe the positive effect of coupling interaction in narrative visualization comprehension and recall. One possible reason that we did not see the benefit for comprehension and recall is that participants did not devote enough time to fully explore both the narratives and visualizations. Another reason could be coupling interaction may distract from the author's intended message. However, we found coupling interaction significantly increased participants' interaction with the story content. Besides, most participants expressed their preference for coupling interaction. For example, a participant wrote, *"I wish all news articles were interactive like this. It makes it so easy to be able to see the part of the chart that the article is discussing or to see the part of the article that the chart is referencing."*

In our study, we also explored how different story layouts affect the reading experience. We adapted a New York Times

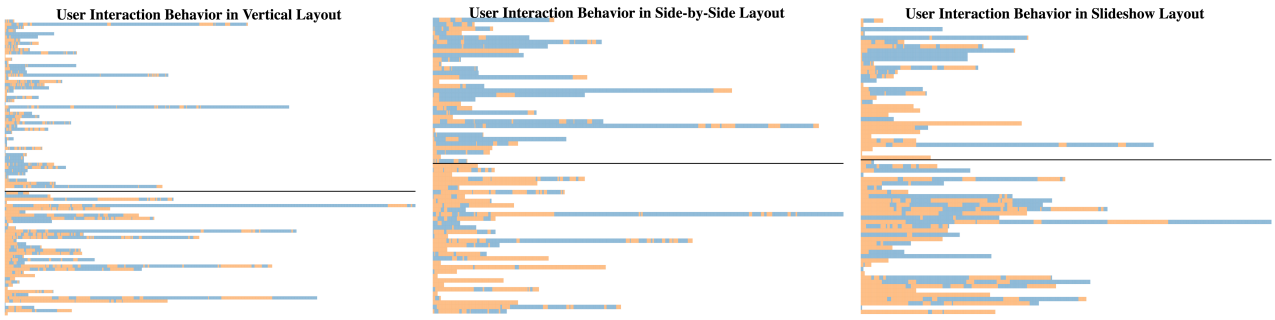

Figure 8. User interaction behavior in different layout settings. The temporal interaction behavior data for each participant is encoded as a row of squares, in which blue square stands for hovering over a visualization element and orange represents hovering over a narrative text. Squares in each row are sorted by time from left to right. The total number of performed interactions varies from one participant to the next, which causes the varying length of rows. Each graph is also separated by a dashed line. Interaction behavior data performed in conditions with coupling interaction is shown below the dashed line.

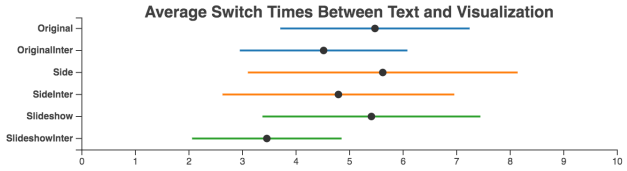

Figure 9. Switch times of hover over interaction between narrative text and visualizations. Error bars are 95% CIs

visual stories with vertical layout into two other different layout settings: side-by-side (with all narrative text on the left and visualizations on the right) and slideshow (with each section of the story as an individual slide and for each slide, narrative text and visualizations are split into two sides). We did not find any evidence to show either of the three layouts is significantly better than others in terms of comprehension, recall, engagement and subjective metrics including confidence and perceived easiness for the given tasks. Nevertheless, we noticed an apparent preference for slideshow layout from the feedback we collected. As a participant said, *" I really like the layout of this study. Very easy to navigate and clear directions."*

Additionally, we analyzed which of the three layout settings is the most appropriate one to add coupling interaction. Our finding show adding coupling interaction can significantly increase user interaction in the vertical and side-by-side layout. This is not surprising because, for the slideshow layout, narrative and visualization are already *visually linked* to each other in the slide format. On the contrary, the side-by-side layout provides a loose mapping between the text and visualizations. We expected coupling interaction to more explicitly link the visualization to its corresponding narrative text and provide benefits for story comprehension. But we did not observe any evidence for such benefits. A possible reason could be the distraction. In the side-by-side layout, when a participant hovers over an element of a visualization, the narrative section will auto-scroll to the corresponding narrative section. This might cause distraction because participants were not familiar with the interaction mode. As one participant stated, *"I enjoyed the live data highlights that showed up as I read. It was a bit distracting but I think that is just because I am new to this format."*

Finally, we focused our analysis to explore the area noted by Segal et al. [32] described as "A promising direction for future research is to focus squarely on readers' experience when viewing and interacting with narrative visualizations." We collected and analyzed user interaction behavior patterns across all conditions. We found participants tended to focused on interacting with visualizations in side-by-side layout. Adding coupling interaction significantly increased the interaction with the text. This result shows the promise of adding coupling interaction on text to engage participants exploring the narrative visualization. We also found that participants switched more frequently in the vertical layout. This makes sense because in the vertical layout participants read individual sections where each includes both narrative text and visualizations from top to bottom.

DESIGN SPACE FOR COUPLING INTERACTION

Our study illustrates the potential benefits and pitfalls of using coupling interaction in narrative visualization. However, the design space for integrating coupling interaction into narrative visualization remains under explored. In the following, we suggest two key design strategies for using coupling interaction to enhance narrative visualization.

Constrained Interaction Settings. The most well-known side effect of interaction is distraction. We expected that coupling interaction would improve the story comprehension by linking the related information in narratives and visualization, however we failed to find evidence for this. We suspect that this is due to the coupling interaction settings in our study causing too much distraction. For example, the auto-scrolling feature in the side-by-side layout may be distracting when participants were exploring the visualization and may also cause disorientation when they return to the story text. Studies showed data stories appear to be most effective when they have constrained interaction [32]. Thus, constrained interaction settings should be considered when using coupling interaction in visual stories.

Advanced Interaction Features. Interaction features, such as zooming, filtering, and searching, help reader explore visu-

alizations more effectively [21]. However, people do not tend to interact with visualizations much. New York Times editor Gregor Aisch noted in his talk, *"I put a switch button at the top of the graph, it can be clicked to show the other half of the graph, but only 10% of the readers click it....I made it bigger and added a conspicuous background color, it only attracted 7% more people.."* To solve this problem, designers may use coupling interaction to perform advanced interaction from the narrative text. For example, in the story *"2015 NCAA Tournament bracket: can you hack it?"* [7], the first graph offers several buttons for filtering different rounds. An alternative option could use coupling interaction to add hover over function on words such as *"Elite Eight"* and *"Final Four"* to filter the graph when readers hover over one of the words, priming the user for these potential interaction modes.

LIMITATION AND FUTURE WORK

Coupling interaction is a technique that explicitly links narrative text with its explanatory visualization element and vice versa. In our study setting, we implemented coupling interaction by manually labeling the relations between narrative text and visualization elements. How can we automatically create the linking? A promising direction for future research is to focus solely on linking narrative text and visualizations. Natural language processing techniques such as text summarization [26] can be employed to extract key ideas from the narratives. There are also tools including Quill [4] that can transform visualizations into human-sounding narratives. Further research is needed to automatically link the narrative and visualization.

Although our study found several significant results, we failed to find evidence that coupling interaction can improve story comprehension. We suspect some comprehension tasks might be too hard to get right. For example, the response accuracy of the **TextVis** comprehension task was only 0.2 - 0.4 out of 1. Furthermore, while our study used a large sample (n=247) to test the six conditions, more participants and additional objective and subjective performance metrics can make the findings stronger.

Since the narrative visualization we used in the paper has several individual sections where each section has its narrative and visualization, we deliberately chose the three different layout settings that are suitable for this story to test the effect on the reading experience. Future work should also evaluate the effect of other layout settings.

Our user interaction behavior analysis identified underexplored regions of viewing and interacting behavior on narrative visualization. Further research is needed to characterize how readers view the narrative visualization for each *genre* of narrative visualization.

REFERENCES

1. Mohammad Allahbakhsh, Boualem Benatallah, Aleksandar Ignjatovic, Hamid Reza Motahari-Nezhad, Elisa Bertino, and Schahram Dustdar. 2013. Quality control in crowdsourcing systems: Issues and directions. *IEEE Internet Computing* 17, 2 (2013), 76–81.
2. Robert Amar, James Eagan, and John Stasko. 2005. Low-level components of analytic activity in information visualization. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on*. IEEE, 111–117.
3. Benjamin Bach, Natalie Kerracher, Kyle Wm Hall, Sheelagh Carpendale, Jessie Kennedy, and Nathalie Henry Riche. 2016. Telling stories about dynamic networks with graph comics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 3670–3682.
4. Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2573–2582.
5. Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2014. Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823* (2014).
6. Jeremy Boy, Francoise Detienne, and Jean-Daniel Fekete. 2015. Storytelling in information visualizations: Does it engage users to explore data?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 1449–1458.
7. Kenan Davis and Kenton Powell. 2015. 2015 NCAA Tournament bracket: can you hack it? *the Guardians* (2015). `https://www.theguardian.com/sport/ng-interactive/2015/mar/17/2015-ncaa-tournament-basketball-bracket`
8. Lydia DePillis, Kulwant Saluja, and Denise Lu. 2015. A Visual Guide to 75 Years of Major Refugee Crises Around the World. *The Washington Post* (2015). `https://www.washingtonpost.com/graphics/world/historical-migrant-crisis/`
9. Evanthia Dimara, Anastasia Bezerianos, and Pierre Dragicevic. 2017. Narratives in Crowdsourced Evaluation of Visualizations: A Double-Edged Sword?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 5475–5484.
10. Fan Du, Ben Shneiderman, Catherine Plaisant, Sana Malik, and Adam Perer. 2017. Coping with volume and variety in temporal event sequences: Strategies for sharpening analytic focus. *IEEE transactions on visualization and computer graphics* 23, 6 (2017), 1636–1649.
11. FORD FESSENDEN, SERGIO PEÃĞANHA JASMINE C. LEE, and ANJALI SINGHVI. 2017. "Immigrants From Banned Nations: Educated, Mostly Citizens and Found in Every State. *New York Times* (2017). `https://www.nytimes.com/interactive/2017/01/30/us/politics/trump-immigration-ban-demographics.html?mcubz=0`
12. Mario R Garcia, Margaret Mary Stark, and Ed Miller. 1991. *Eyes on the News*. Poynter Institute.

[4] https://narrativescience.com/Platform

13. Nahum Gershon and Ward Page. 2001. What storytelling can do for information visualization. *Commun. ACM* 44, 8 (2001), 31–37.

14. Richard M Golden and David E Rumelhart. 1993. A parallel distributed processing model of story comprehension and recall. *Discourse processes* 16, 3 (1993), 203–237.

15. David Gotz and Harry Stavropoulos. 2014. Decisionflow: Visual analytics for high-dimensional temporal event sequence data. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 1783–1792.

16. Jeffrey Heer and Ben Shneiderman. 2012. Interactive dynamics for visual analysis. *Queue* 10, 2 (2012), 30.

17. Kenneth Holmqvist, Jana Holsanova, Maria Barthelson, and Daniel Lundqvist. 2003. Reading or scanning? A study of newspaper and net paper reading. *Mind* 2, 3 (2003), 4.

18. Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240.

19. Jessica Hullman, Nicholas Diakopoulos, and Eytan Adar. 2013a. Contextifier: automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2707–2716.

20. Jessica Hullman, Steven Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. 2013b. A deeper understanding of sequence in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2406–2415.

21. Robert Kosara and Jock Mackinlay. 2013. Storytelling: The next step for visualization. *Computer* 46, 5 (2013), 44–50.

22. Heidi Lam. 2008. A framework of interaction costs in information visualization. *IEEE transactions on visualization and computer graphics* 14, 6 (2008).

23. Bongshin Lee, Rubaiat Habib Kazi, and Greg Smith. 2013. SketchStory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2416–2425.

24. Kwan-Liu Ma, Isaac Liao, Jennifer Frazier, Helwig Hauser, and Helen-Nicole Kostis. 2012. Scientific storytelling using visualization. *IEEE Computer Graphics and Applications* 32, 1 (2012), 12–19.

25. Narges Mahyar, Sung-Hee Kim, and Bum Chul Kwon. 2015. Towards a taxonomy for evaluating user engagement in information visualization. In *Workshop on Personal Visualization: Exploring Everyday Life*, Vol. 3. 2.

26. Inderjeet Mani and Mark T Maybury. 1999. *Advances in automatic text summarization*. Vol. 293. MIT Press.

27. Radan Martinec and Andrew Salway. 2005. A system for image–text relations in new (and old) media. *Visual communication* 4, 3 (2005), 337–371.

28. S McKenna, N Henry Riche, B Lee, J Boy, and M Meyer. 2017. Visual Narrative Flow: Exploring Factors Shaping Data Visualization Story Reading Experiences. In *Computer Graphics Forum*, Vol. 36. Wiley Online Library, 377–387.

29. Heather L O'Brien and Elaine G Toms. 2010. The development and evaluation of a survey to measure user engagement. *Journal of the Association for Information Science and Technology* 61, 1 (2010), 50–69.

30. Bahador Saket, Alex Endert, and John T Stasko. 2016. Beyond Usability and Performance: A Review of User Experience-focused Evaluations in Visualization.. In *BELIV*. 133–142.

31. Arvind Satyanarayan and Jeffrey Heer. 2014. Authoring narrative visualizations with ellipsis. In *Computer Graphics Forum*, Vol. 33. Wiley Online Library, 361–370.

32. Edward Segel and Jeffrey Heer. 2010. Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148.

33. R Core Team. 2014. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. (2014).

34. Christian Tominski. 2015. Interaction for visualization. *Synthesis Lectures on Visualization* 3, 1 (2015), 1–107.

35. Krist Wongsuphasawat, John Alexis Guerra Gómez, Catherine Plaisant, Taowei David Wang, Meirav Taieb-Maimon, and Ben Shneiderman. 2011. LifeFlow: visualizing an overview of event sequences. In *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM, 1747–1756.

36. Ji Soo Yi, Youn ah Kang, and John Stasko. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE transactions on visualization and computer graphics* 13, 6 (2007), 1224–1231.