

# Linking and Layout: Exploring the Integration of Text and Visualization in Storytelling

Qiyu Zhi<sup>1</sup>, Alvitta Ottley<sup>2</sup>, & Ronald Metoyer<sup>1</sup>

<sup>1</sup>University of Notre Dame, Notre Dame, IN

<sup>2</sup>Washington University in St. Louis, St. Louis, MO

## Abstract

Modern web technologies are enabling authors to create various forms of text visualization integration for storytelling. This integration may shape the stories' flow and thereby affect the reading experience. In this paper, we seek to understand two text visualization integration forms: (i) different text and visualization spatial arrangements (layout), namely, vertical and slideshow; and (ii) interactive linking of text and visualization (linking). Here, linking refers to a bidirectional interaction mode that explicitly highlights the explanatory visualization element when selecting narrative text and vice versa. Through a crowdsourced study with 180 participants, we measured the effect of layout and linking on the degree to which users engage with the story (user engagement), their understanding of the story content (comprehension), and their ability to recall the story information (recall). We found that participants performed significantly better in comprehension tasks with the slideshow layout. Participant recall was better with the slideshow layout under conditions with linking versus no linking. We also found that linking significantly increased user engagement. Additionally, linking and the slideshow layout were preferred by the participants. We also explored user reading behaviors with different conditions.

## CCS Concepts

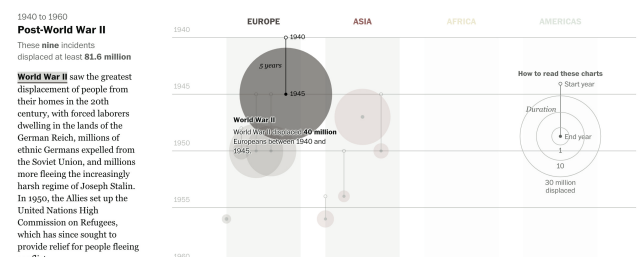
• **Human-centered computing** → **User studies; Visualization design and evaluation methods; Web-based interaction;**

## 1. Introduction

Narrative and visualizations are increasingly being used together to reveal and convey data-rich stories. News organizations and scientific platforms often incorporate data visualization as a tool for understanding alongside their written content [SH10]. Analysts embed graphics into their reports to better communicate their findings [KM13]. Modern web technologies (e.g. D3 [BOH11] and Vega-lite [SMWH17]) have further facilitated this growth in popularity and enabled the efficient creation of a wide range of interactive visual stories with innovative techniques.

This technology development has fostered a variety of text and visualization integration forms. For example, as shown in Figure 1, the text and visualization are placed side by side and can be interactively linked by clicking underlined words. In “The World According to China” [AKR15], the text and visualization are separated in different views and can be switched by scrolling. The proliferation of *Storytelling or Narrative Visualization* has also led the visualization community to develop analytic frameworks to inform design strategies [HD11, SH10], study how narrative flows [MHRL\*17] and storytelling [BDF15] affect the story reading experience, and implement tools [HDA13, SH14, FBM16] for the creation of narrative visualizations. While we have made significant progress to-

ward understanding the utility of storytelling, little attention has been given to *how* we should integrate text and visualization.



**Figure 1:** An example story using a linked text and visualization technique. When readers click “World War II”, the corresponding circle will be highlighted to show the related information

In this paper, we focus on two ways of integrating narrative text and visualizations: interactively linking (linking) and spatially arranging (layout). Segel and Heer present a design space consisting of *genres* or types of narrative visualizations [SH10]. These *genres* are different approaches to communicating the visual narratives.

One of the key variables in these *genres* is in how text and the visual elements are spatially arranged to enable effective communication of the story. We refer to this spatial arrangement of text and visualization as **layout** and study how different layout settings affect the story reading experience in terms of comprehension, recall, and user engagement with the story.

Another approach to integrating text and visualization is through interactively linking the explanatory text with its visual counterpart. For example, “A visual guide to 75 years of major refugee crises around the world” [DSL15] from the Washington Post tells a story about major refugee crises around the world by explicitly linking text and the corresponding visual elements (see Figure 1). This technique of **linking of elements through interaction** has been used in stories such as *Visualizing MBTA Data* [BC13], *The Race Gap in American Police Departments* [AP15] and *Road Map* [AM15]. While linking has been used in these recent stories, there is a lack of understanding of how it might affect the story reading experience. In this paper, we use the term **linking** to refer to a *bidirectional* interaction mode that explicitly highlights the explanatory visualization element when selecting narrative text and vice versa.

We designed and conducted a study to measure if adding linking or changing story layout can affect measures such as comprehension, recall, user engagement, and story reading behaviors. We applied linking to an existing visual story and also adapted the story into two common layouts: vertical and slideshow. We then launched the study in Amazon Mechanical Turk with 180 participants. Our results indicate that linking can significantly improve user engagement. The results also show the benefit of the slideshow layout on comprehension and recall tasks compared to the vertical-layout conditions. We explore the difference in reading behaviors among different conditions and discuss integrating text and visualizations in visual story design.

## 2. Related Work

### 2.1. Visual Storytelling and Narrative Visualization

Technology provides us with new media and genres that we can now use to convey information in a story-like fashion [GP01]. Narrative visualization that integrates storytelling into information visualization is a popular research area in the visualization community [KM13]. Segel and Heer 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 [SH10]. Hullman and Diakopoulos also proposed an analytic framework “*visualization rhetoric*” to understand how design techniques for “tell a story” can affect a user’s interpretation [HD11].

Researchers have proposed multiple tools and conducted various studies to support authoring and understanding narrative visualizations. *SketchStory*, for example, uses “real-time freeform sketching” to support authors to create personalized and expressive stories on a whiteboard [LKS13]. *Ellipsis* combines a graphical interface with a “domain-specific language for storytelling” to enable journalists to create interactive narratives [SH14]. *Contextifier* automatically generates an annotated stock chart given some news

text [HDA13]. To understand the effect of various design factors in narrative visualizations, Boy et al. evaluated how initial narratives in exploratory visualizations affect user engagement [BDF15]. Kong et al. studied the effect of “visual cues” on users attention [KLK17]. Layout is also an important factor that determines the way in which text and graphs are spatially organized in the story. The seven genres characterized by Segel and Heer vary in terms of ordering of visual elements, which implies differences in layout. McKenna et al. systematically investigated and studied the “flow factors” in narrative visualization and explored their effect on user engagement [MHRL\*17].

In contrast to previous studies, we focus on the integration forms of text and visualization in narrative visualizations and seek to understand how linking and spatial layout settings impact the story reading experience.

### 2.2. Interactively Linking Text and Visualization in Storytelling

Cognitive science researchers have studied how people comprehend multimedia content for many years. Hegarty and Just argued that learning can be seen as having three kinds of cognitive demands: *essential processing*, *incidental processing*, and *representational holding* [HJ89]. The associated *representational holding* can be potentially reduced and *essential processing* can be increased if readers can easily link the text and pictures in multimedia content. Mayer presented the multimedia principle, which implies the integration of text and picture information can help individuals learn better than from text information alone [May02]. To provide evidence of this theoretical assumption, Mason et al. used the eye-tracking methodology to examine how fourth graders process text and graphics and confirmed a greater text-figure integrative processing was associated with higher learning performances [MTP13].

In the context of visualization research, multiple coordinated visualizations enable users to explore information through *brushing and linking* and *overview and detail views* [NS00]. Similarly, making connections between narrative visualization elements could be also useful to understand the story particularly when data represented in visualizations serves as an exploratory role. For example, in the New York Times story “Steroids or Not, the Pursuit is On”, the three main sections are graphically linked together by color and shape. The linking helps readers quickly discern the reference between the image and other annotations [SH10]. Through an analysis of recent data-driven stories, Stolper et al. summarized four categories that help authors tell stories in creative ways: communicating narrative and explaining data, linking separated story elements, enhancing structure and navigation, and providing controlled exploration [SLRS17]. In this paper, we focus specifically on linking narrative text and visualization elements.

Text and visualization can both play an explanatory role for each other in narrative visualizations. Yet, identifying references between text and visualization can be challenging. Segel and Heer also noted that current news visualization seems to lack narratives that highlight key findings in the visualization [SH10]. To address this issue, Kwon et al. developed *VisJockey* that provides a new

way to read stories by connecting text and visualizations [KSJ\*14]. Kong et al. designed a crowdsourcing pipeline to extract the text and visualization references [KHA14]. Metoyer et al. presented an approach to automatically “couple narrative text to visualizations for data-rich stories.” [MZJS18] by extracting data-centric references from the text and linking them to the visualization.

In this paper, we adapted a linking technique into an existing visual story where the explanatory visual element is highlighted when users mouse over narrative text, and vice versa. We then evaluate this linking effect on reading visual stories.

### 2.3. Evaluating Reading Experience in Storytelling

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 [KM13]. To investigate the effect of adding embellishments into visualization, Bateman et al. measured participants’ interpretation accuracy and long-term recall by asking description questions [BMG\*10]. Multiple studies also explored the memorability of visualizations [BVB\*13, SSB15]. Dimara et al. specifically explored the effects of providing task context when evaluating visualization tools using crowdsourcing [DBD17]. They evaluate the comprehension and recall by measuring participants’ ability to “perform and understand” several tasks. We adapted these methods in our study and design specific tasks to measure participants’ comprehension and recall.

Engagement is also an important but complex factor in determining the effectiveness and impact of narrative visualizations as it lacks a unified definition [MKK15]. As such, appropriate “proxies” are needed to describe users’ involvement. McKenna et al. adopted a subjective questionnaire for measuring engagement in different visual story experiment settings [MHRL\*17]. However, the results can be heavily affected by individual differences. Boy et al. evaluated engagement by analyzing user-semantic operations and depth of interaction [BDF15]. Time spent on subjective reaction cards for capturing user feelings were also measured to assess user experience [SES16].

Given one of our goals is to see whether adding linking or changing layout can help engage users in the visual story exploration, we evaluate engagement through a combination of reader’s subjective feedback and investment in the exploration of a narrative visualization. We used the 14 subjective questions that originated from O’Brien et al. [OT10] and tailored by McKenna et al. [MHRL\*17]. Additionally, we recorded the time spent and depth of interaction, which is interpreted as the number of interactions a user performs [BDF15], as an objective metric of engagement.

### 3. Research Questions

We framed three research questions to guide our investigation into the effect of linking and layout on reading experience. We operationalized the reading experience measurement with engagement metrics (user-perceived, time spent, and interaction in-depth), comprehension tasks performance, recall tasks performance, and subjective feedback. We also examined users’ story reading behavior.

#### RQ1: Can linking improve story reading experience?

The main purpose of this study was to investigate the effect of integrating (bi-directional) linking into an existing established narrative visualization.

#### RQ2: How do different layout settings affect reading experience?

Specifically, the narrative visualization used in this paper contains several independent sections where each section includes narrative text and visualizations. Our study conditions included this layout (vertical) as well as the slideshow layout (see Figure 3). We aimed to identify whether changing the layout setting impacted the reading experience.

#### RQ3: How does adding linking or changing layout impact users’ story reading behaviors?

We were also interested in how would users’ behavior changes when adding linking or varying the layout. The appropriate quantitative proxies that can capture reading behaviors include transitions between text and visualization or reading navigation traces.

## 4. Crowdsourced Study Design

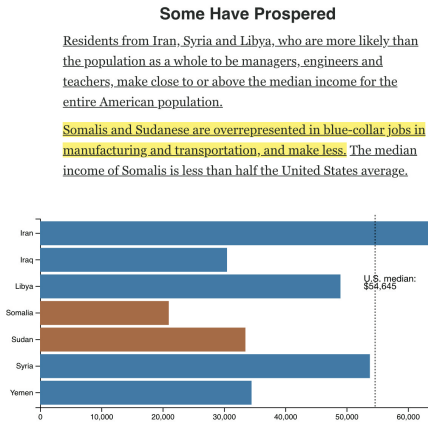
We conducted a large scale between-subjects crowdsourced study using Amazon Mechanical Turk (AMT) to understand how linking and different layout settings affect the visual stories’ reading experience.

### 4.1. Study Condition Settings

We utilized the story “*Immigrants From Banned Nations: Educated, Mostly Citizens and Found in Every State*” by Ford Fessenden et al. of the New York Times [FJSA17]. We selected this story because it had many separate sections with each section including both narrative text 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 narrative text and visualizations of the story.

The original story is presented as a vertical layout where the story can be scrolled up and down (as shown in Figure 2). To explore the effect of changing layout settings on the visual story reading experience, we adapted the original story into another layout setting: slideshow (as shown in Figure 3). We selected vertical (maps to “magazine style”) and slideshow layouts because they are two of the seven basic *genres* that are commonly used in narrative visualization design [SH10].

Additionally, to understand the effects of linking, we created conditions with and without linking for each layout setting. In conditions with linking, an underline indicates the sentence is “hoverable”. When a participant hovered over an underlined sentence, as Figure 2 shows, the explanatory visualization elements related to this sentence were 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.



**Figure 2:** The study interfaces for original with linking study condition. It shows that the explanatory visualization elements are highlighted when participants hovering over the first sentence of the second paragraph in section “Some Have Prospered”.

In total, we have two layout settings and each layout setting has two variations of linking resulting in four conditions:

- **Original:** This condition is our baseline condition where visualizations are individually embedded into narrative text with a vertical layout setting (Figure 2).
- **Original with linking:** This condition is identical to the Original condition but with linking integrated (Figure 2).
- **Slideshow:** In this condition, each section of the narrative visualization is presented as an individual slide with narrative text on the left and visualizations on the right (Figure 3). The buttons at the bottom can be clicked to navigate between slides.
- **Slideshow with linking:** Identical to the Slideshow condition with linking integrated (Figure 3).

## 4.2. Measures

We quantified reading experience by measuring comprehension, recall, engagement, subjective feedback, and reading behavior.

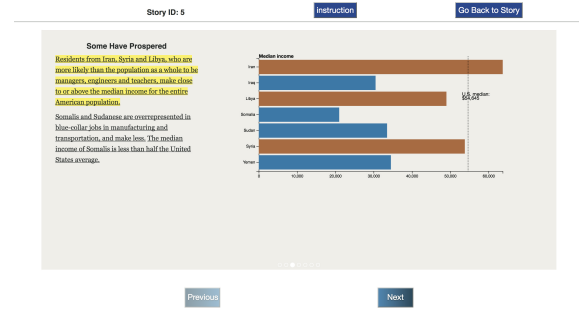
### 4.2.1. Comprehension

Our considerations for comprehension evaluation were two-fold. First, narrative visualizations can be regarded as ordered sequences of narrative text and visualizations, therefore, participants needed to fully comprehend not only text but visualizations as well. The designed tasks include:

**Text (Text):** task which can be completed solely by reading text in the narrative visualization.

**Visualization (Vis):** task which can be completely solely by observing the visualizations.

**Text + Visualization (TextVis):** task which can be completed only by combining an understanding of both the text and visualizations. For example, in the TextVis question in Table 1, participants can only answer it by first finding the two countries in the narrative text and then finding the year in the corresponding visualization.



**Figure 3:** The study interfaces for slideshow with linking study condition. It shows that the explanatory visualization elements are highlighted when participants hovering over the first sentence in section “Some Have Prospered”.

Second, we needed to evaluate if readers could make informed decisions by retrieving relevant information from the narrative visualization [KM13]. Similar to the study of narratives in crowd-sourced evaluation [DBD17], we adapted two taxonomies of low-level information retrieval tasks [AES05] for visualization tasks:

**Extremum (Ext):** task where participants had to find the extreme value from visualizations (e.g. find the largest value in a bar chart).

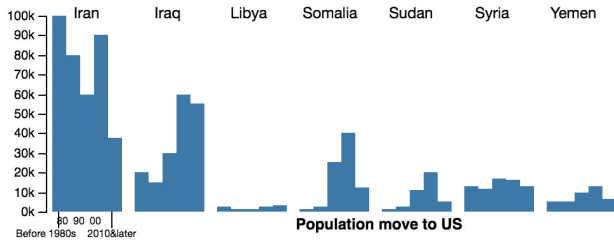
**Comparison (Com):** task where participants had to compare visualization elements across the chart (e.g. compare two values in a bar chart).

| task           | question                                                                                                                                                                    |
|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Vis<br>Ext     | Of the countries mentioned in the article, which country has the highest immigration population in the US?                                                                  |
| TextVis<br>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? |
| Text<br>Com    | Which of the following cities has the most Somali-born residents?                                                                                                           |

**Table 1:** Three example questions. The first question is a **Vis** and **Ext** task, the second is a **TextVis** and **Com** task, the third is a **Text** and **Com** task.

For instance, the visualization in Figure 4 shows the trend of immigrants from the seven countries moving into the US. The first question in Table 1 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 a **Ext** task. For the second question, participants must first determine in “which two countries mentioned in the article were the immigrants better educated than the rest of America” from the text, then they must determine “which year accounts for the largest population move to the US”. Thus this question is a **TextVis** task. In addition, it also





**Figure 4:** The “population move to US” chart in the narrative visualization. It shows the arrival pattern of immigrants from the seven countries.

required participants to compare the values, 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.

#### 4.2.2. Recall

We also measured participants’ ability to recall key points of the narrative visualization. Similar to the comprehension task design, we considered both text and visualizations. Specifically, our designed recall tasks can be summarized as:

**Overview** (Overview): task where participants had to remember the overall idea and select an aspect that had not been discussed in the story.

**Detail** (Detail): task where participants had to remember specific details such as which country had not been mentioned in the story.

**Visualization** (Vis): task where participants had to identify which chart was presented to them in the story.

In total, we used three recall tasks and assigned a binary score of 1 for the correct answer and 0 for all other answers.

#### 4.2.3. Engagement

As discussed in section 2.3, our methods to measure engagement were twofold. First, we adopted the 14 subjective questions to measure user-perceived engagement where each question received a 5-point Likert response from every participant. Second, we measured users’ investment on the story reading by recording the time spent and depth of interaction, which we defined as the number of “hover over” interactions a user performs on the narrative visualization.

#### 4.2.4. Subjective Metrics

We also used subjective metrics to collect individual preferences and to complement objective metrics by asking participants feedback pertaining to the study, layout setting, and linking technique.

#### 4.2.5. Reading Behavior

Our third research question concerns how linking and layout affect users’ reading behavior. However, verifying such a qualitative

question in a web-based crowdsourcing experiment is difficult. It should be noted that we recorded the interaction trace for our participants so we knew the time and location of each hover-over action. Intuitively, linking helps readers make connections between text and visualization elements, especially when the visualizations are used to explain aspects of the narratives. Additionally, Heer and Robertson studied that linking elements through animation has a strong effect on “retaining readers in the context” [HR07]. Inspired by this work and based on the collected data, we decided to operationalize the reading behaviors with two quantitative metrics:

- **Transition:** Transition measures how many times the user transitions between text and visualization. Note that transition includes transition from text to visualization (Text-to-Vis) as well as from visualization to text (Vis-to-Text).
- **Confusion:** As mentioned before, the story we selected consists of several independent sections where each includes a text paragraph and a visual chart. We noticed participants in pilot test sometimes mapped a nearby but wrong visualization to the paragraph they were reading under vertical-layout conditions. To measure this effect, we define confusion as an event where a user conducted a series of hover over actions (more than three) on a section, then hovers over an incorrect adjacent section for once or twice, and went back to the previous section. Note that the confusion only happens in conditions with the original vertical layout.

#### 4.3. Pilot test

To avoid unexpected problems in crowdsourcing study, we first recruited four participants in our lab to read through the immigrants story under the four conditions, respectively. Upon finishing, participants were asked to solve comprehension and recall tasks and rate the engagement questions.

All the participants agreed the story was easy to follow and comprehension and recall tasks were solvable according to the story content. Three participants went back to the story at least once when answering comprehension tasks. Two participants pointed out that the time interval between comprehension and recall tasks should be longer. We also noticed some accidental “hover over” events in other conditions after analyzing their interaction data.

Based on the pilot test, we decided to (i) put recall tasks after engagement questions, (ii) configure mouse sensitivity and the time interval to avoid the accidental “hover over” events.

#### 4.4. Participants

We recruited 180 participants from AMT where each participant 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. The participant was offered a maximum \$2.20 reward. Participants had a varied education background: 10% had a master’s or advanced degree, 55.3% had a bachelor’s degree, 34.7% had high school or some college experience, 48% were female, 52% were male, nine participants were immigrants, and only 4% were previously familiar with the subject matter presented in the story.

#### 4.5. Procedures

The study followed a between-subjects design. Each participant received the narrative visualization with a random condition. After reading an introduction of the study, participants were given a tutorial as a guided tour through the study interface. The tutorial was implemented as multiple tooltips. For participants in conditions with linking, we showed them the effect of hovering over narrative text and visualizations. For participants in conditions with the slideshow layout setting, we showed them how to navigate the narrative visualization by clicking the buttons.

Participants were asked to carefully read and fully explore the story no matter how long it would take. After reviewing all narrative text 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.

Like prior studies of narrative visualization comprehension, we placed the task questions after the story rather than within the story [DBD17]. Participants were allowed to 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 3) when answering comprehension tasks. The button was disabled when they were answering the recall questions. After submission, participants were given a random id to paste into AMT to complete their HIT.

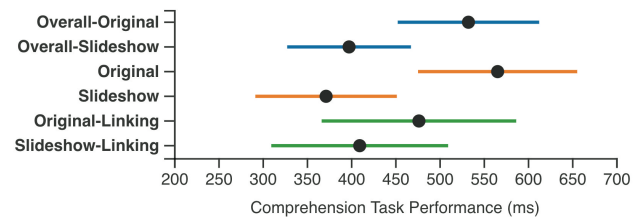
#### 4.6. Hypotheses

Based on prior work, we formulated hypotheses to answer our research questions.

The linking interaction mechanism can be seen as a method for reducing the associated *representational holding* and enhancing the *essential processing* of readers as they read and comprehend a narrative visualization. Hegarty and Just [HJ89] showed that readers tend to read the text first and then look at the corresponding graphics. Similarly, in narrative visualizations, where the text and graphics are separate, readers may first read the textual story and then visually scan the whole visualization in search of the corresponding information. Hence we hypothesize participants in linking conditions perform better in comprehension and recall tasks and engage with the story more deeply.

Our second research question concerned the effect of layout settings. Given that the slideshow structure allows users to walk through the story step by step and provides a clear mapping between text and visualization, we expected the participants in slideshow-layout conditions will exceed participants in the vertical layout condition in terms of comprehension and recall tasks.

Regarding the effect on reading behaviors, given the clear text visualization mapping in the slideshow layout or linking conditions, we expected there to be fewer text visualization transition actions under these conditions.



**Figure 5:** The Comprehension IES for all participants across the different conditions in the study. Higher scores indicates lower performance and the error bars are 95% CIs. We found that the slideshow is more effective than the vertical layout at supporting comprehension.

### 5. Results

We excluded two participants for participating multiple times and eleven participants for spending extremely short or long times on reading the story or answering questions. As a result, we accepted and analyzed a total of 167 jobs. For each participant, we recorded a 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, 14 5-point Likert responses for engagement statements, three recall task responses, the time spent (including story reading time, questionnaire completion time, and time for completing comprehension tasks, recall tasks, demographic questions, and engagement tasks), reading behaviors including each text hover over event (including timestamp and text content) and each visualization element hover event (including timestamp and visualization element ID). We mainly used the unpaired two-sample t-test to compare task performance between conditions. The R code and detailed analyses can be found in the **Supplemental Materials**.

#### 5.1. Comprehension Tasks

Inspired by previous studies [CCH\*14], to measure the effect of linking and layout on comprehension tasks, we look at both task completion time and task accuracy as performance measures. The comprehension tasks completion time was normally distributed ( $M = 256, SD = 33$ ). Considering participants may focus on generating the correct answer while sacrificing time, we used a combined score for task performance, known as the Inverse Efficiency Score (IES) [TA83]. For each participant, the IES is the task completion time divided by the proportion of correct responses. The performance result is reported in seconds and a higher score indicates a lower performance.

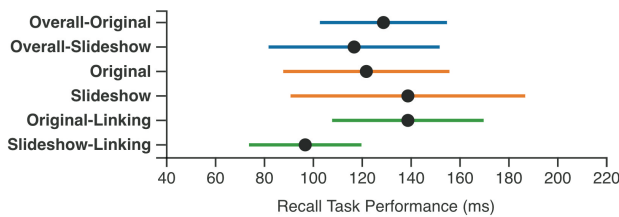
We did not observe major differences between linking and no-linking conditions on the comprehension task performance. However, we found layout significantly affected comprehension tasks performance, as shown in Figure 5. The overall result covers both linking and no-linking condition. We observed a significantly better ( $t = 3.6, p < 0.01$ ) task performance in the slideshow layout ( $M = 533, SD = 81$ ) than in the original vertical layout ( $M = 398, SD = 72$ ). Specifically, under no-linking conditions, partici-

pants performed significantly better in the slideshow layout ( $t = 2.6, p < 0.01$ ). The difference becomes moderate under linking conditions ( $t = 2.1, p = 0.014$ ).

As such, we found the slideshow layout is more effective than the vertical layout at supporting participants' comprehension.

## 5.2. Recall Task

The recall tasks completion times were normally distributed ( $M = 86, SD = 9.89$ ). We also used Inverse Efficiency Score (IES) to measure the recall task performance. Again, a higher score stands for a lower performance.



**Figure 6:** The Recall IES for all participants across the different conditions in the study. Higher scores indicates lower performance and the error bars are 95% CIs. We found that slideshow is more effective than vertical layout for recall tasks under linking conditions.

We did not find any evidence for linking improving the recall task performance. The layout's effect on recall task performance results are shown in Figure 6. Contrary to our expectation, the overall performance of recall tasks did not significantly differ in the original vertical layout and the slideshow layout ( $t = 1.2, p = 0.2$ ). However, we observed a significant difference ( $t = 2.13, p = 0.03$ ) between the original layout ( $M = 97, SD = 23$ ) and the slideshow layout under linking conditions ( $M = 139, SD = 31$ ).

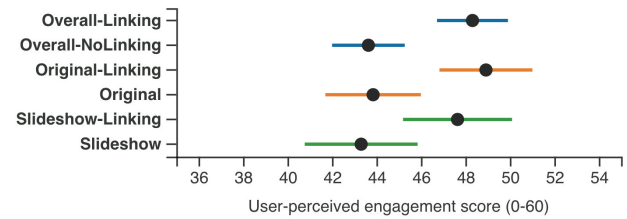
Thus, we observed under linking conditions, slideshow is more effective than the vertical layout for recall tasks.

## 5.3. Engagement

Regarding engagement measurement, we report questionnaire results, user's investment in the story (time spent), and the *depth of interaction* (hover interaction number).

### 5.3.1. User-perceived engagement

We did not find any major differences between conditions with the original vertical layout and the slideshow layout for user-perceived engagement results. In Figure 7, we show the result of linking effect on user-perceived engagement questionnaire scores. The overall result covers both the original vertical layout and the slideshow layout. We observed a significantly better ( $t = -4, p < 0.001$ ) engagement score in linking condition ( $M = 48.32, SD = 1.59$ ) than in no-linking condition ( $M = 43.64, SD = 1.63$ ). Specifically, under vertical-layout conditions, participants engaged significantly more

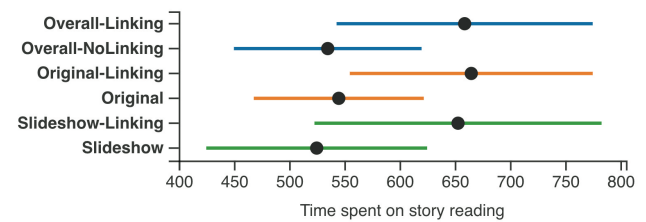


**Figure 7:** User-perceived engagement results for linking and no-linking conditions. Error bars are 95% CIs. Results show participants in linking conditions had higher engagement.

in linking condition ( $t = -3.3, p = 0.001$ ). The difference is also significant under the slideshow layout ( $t = -2.4, p < 0.01$ ).

In summary, we found linking improved user-perceived engagement.

### 5.3.2. Time spent on story reading



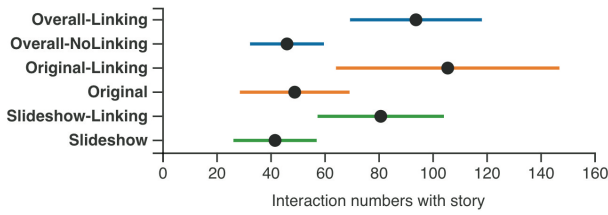
**Figure 8:** Time spent results for linking and no-linking conditions. Error bars are 95% CIs. Results show participants spent more time on reading the story under linking conditions.

We compared time spent on reading the story as a proxy to examine the effect of linking or layout on users' investment in the story. There is no evidence that users' story reading time differs for various layout settings. However, we found participants in conditions with linking spent significant more time than in no-linking conditions (Mean: Linking=659.24, NoLinking=539.04,  $t = -2.05, p = 0.03$ ) (Fig 8). The result also stands in vertical-layout conditions ( $t = -2.35, p = 0.026$ ) and slideshow conditions ( $t = -1.78, p = 0.042$ ).

As such, linking improved users' time investment on story reading.

### 5.3.3. Interaction numbers

Based on previous study, we use the recorded hover-over interaction numbers as an proxy to evaluate users' depth of interaction. We did not find any major differences between conditions with the original vertical layout and the slideshow layout. We observed that participants in conditions with linking interacted significantly more with the story than those in no-linking conditions (Mean:



**Figure 9:** User-performed interaction results for linking and no-linking conditions. Error bars are 95% CIs. Results show participants performed more interactions under linking conditions.

Linking=93.95, NoLinking=46.02, Fig 9,  $t = -3.34, p = 0.01$ ). The significance was also found in vertical-layout conditions ( $t = -2.45, p = 0.02$ ) and slideshow conditions ( $t = -2.73, p < 0.01$ ), respectively.

As a result, linking improved participants' depth of interaction on story reading.

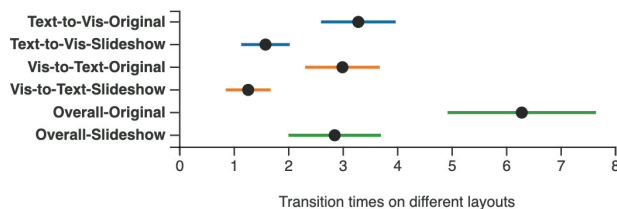
#### 5.4. Subjective Feedback

We collected user feedback to gather insights on their subjective preferences for linking and different layout settings. As we expected, the slideshow layout was favored by most participants, "I really like the layout of this study. Very easy to navigate and clear directions." Our participants also thought highly of the potential of linking. 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." We also saw a few concerns regarding the use of linking: "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."

#### 5.5. Reading Behaviors Observation

We now investigate how linking and layout impacts reading behavior.

##### 5.5.1. Transition



**Figure 10:** User-performed transition results for various layout setting conditions. Error bars are 95% CIs. We found participants performed more Text-to-Vis transitions in the vertical layout.

Figure 10 shows the different transition results over the original vertical layout and the slideshow layout settings. We observed participants transitioned significantly more in the vertical layout, including Text-to-Vis transition ( $t = 4.08, p < 0.001$ ), Vis-to-Text transition ( $t = 4.24, p < 0.001$ ), and overall transition ( $t = 4.19, p < 0.001$ ). The significance also stands in both linking and no-linking conditions.

Regards to linking effect on transition, we only observed a moderate significance in slideshow-layout conditions ( $t = -2.2, p = 0.03$ ), where participants in linking condition ( $M = 3.73, SD = 1.3$ ) transitioned more than in no-linking condition ( $M = 1.9, SD = 0.98$ ).

#### 5.5.2. Confusion

We found the confusion numbers in original condition ( $M = 0.89, SD = 0.3$ ) is significantly higher ( $t = -3.39, p = 0.001$ ) than in original with linking condition ( $M = 0.29, SD = 0.16$ ). Layout did not affect confusion number significantly.

#### 6. Discussion

The results of our study suggest that the layout of a storytelling visualization can significantly impact comprehension when there are multiple visualization components. We found that performance on comprehension tasks was better in the slideshow layout than the vertical layout. The design of the slideshow isolated the different segments of the story and created a one-to-one mapping between each paragraph and its corresponding visualizations. For the vertical layout, this mapping was less clear. In some instances, the layout may have introduced mapping ambiguity when there is text both above and below a visualization. We suspect that the clear text visualization mapping structure of the slideshow is the underlying feature that drives this finding.

We analyzed passive mouse hovering data to understand how participants interacted with the different layout settings and for further insight into the "slideshow advantage". We found that participants performed significantly more Text-to-Vis transition actions in the vertical layout compared to the slideshow layout. This finding provides suggestive evidence that, when using the vertical layout, participants may have struggled to integrate the information across the text and visualization formats. The subjective feedback also seems to corroborate this. Participants preferred the slideshow layout and found it more clear and convenient. However, future work is needed to investigate this effect. Adding linking erased the slideshow advantage. We found no significant performance difference between the slideshow and the vertical layout when we highlighted the text-vis components through linking. We also found linking decreased users' "confusion" actions in the vertical layout significantly. We measure "confusion" by counting how many times participants mismatched the text and the corresponding visual elements. This finding confirms our hypothesis that linking provides a clear reference between text and visualization. Users' feedback also indicated that linking is helpful for finding the corresponding reference in the visualization.

Another benefit to adding linking is it increases engagement.



We observed a significant effect on all of the three metrics: user-perceived ratings, invested time, and depth of interaction. We acknowledge that engagement is a complex, multidimensional construct. No single metric is sufficient. For example, longer story reading time does not necessarily indicate that a participant is engaged. The interaction time is especially noisy for crowdsourced studies because it is impossible to tell if participants were merely distracted. While these concerns may complicate our understanding of user engagement, we propose that the combination of all three metrics supports an argument for improved engagement.

Our study serves as a starting point to explore the various ways in which text and visualizations can be associated and arranged in narrative visualizations and how they affect users' reading experiences. Segel and Heer [SH10] refer to narrative tactics used in other art forms such as film and comics to present their genres of narrative visualizations. Similarly, we can refer to these very related work to explore the interplay of text and visualizations. For example, McCloud presents categories of words and picture combinations in comics and the different effects they produce, such as *word-specific* where words dominate the picture while in *picture-specific* where pictures are dominating and words only provide supporting information [McC11]. These ideas can be adapted to create similar effects in narrative visualizations as well. For example, *word-specific* approaches can be used in author-driven stories where there is an overarching narrative. McCloud also describes the ways in which pictures and words can be arranged in a page and within a panel in comics to produce desired effects [McC11]. This translates to layouts in our discussion. Through our study, we hope to encourage researchers to explore such combinations and arrangements in narrative visualizations, and perform studies such as ours to study their effects on reading experiences.

## 7. Limitation and Future Work

Our study is not without limitations. While we believe that linking techniques can be generalized and applied to any other narrative visualizations, our study results are presently applicable only to the narrative visualization used. It is possible that other factors such as the story topic, limited number of layout settings, and participants' display size may have influenced the study results. We also acknowledge that the study results are limited in a crowdsourcing study setting which may not represent real-world visual story reading situations. For example, organizations such as the *New York Times* have scaled back their creation of interactive visualizations in lieu of static images because users often ignore the interactive component in the story [Tse16]. We found that the slideshow layout is better in terms of comprehension, but it's also possible that users are less likely to get to the end of an article in the slideshow layout compared to the vertical-layout setting. Future work should explore different types of stories, different visualization designs, adapt to more layout settings, and consider the real-world usage situations.

Following existing research [BOZ\*14], we use mouse hover as a proxy for users' selective attention. Although prior work demonstrates a strong correlation between mouse movements and eye gaze interactions [RF07], future work may benefit from measuring eye gaze directly. The eye tracking data may also provide more

definitive insight into whether or not user struggle to integrate text and visualization in the vertical layout. Future work could also consider other interaction techniques such as clicking, zooming, scrolling, and dragging to accurately capture the "user investment" in reading the story.

In our study setting, we implemented linking by manually labeling the relations between narrative text and visualization elements. How can we automatically generate the linking? A promising direction for future research is to focus on the automatic linking of narrative text and visualizations. Kong et al. developed a crowdsourcing pipeline to extract the references between text and visualization [KHA14]. Metoyer et al. coupled sports visualizations with stories by identifying subject and time information [MZJS18]. Natural language processing techniques such as text summarization, named entity identification, or subject-verb-object extraction [MM99] can also be employed to extract key ideas, names, or relations from the narrative text. There are also tools such as Quill that can transform visualizations into human-sounding narrative text. Further research is needed to automatically link the narrative text and visualization in meaningful ways.

## 8. Conclusion

We study two ways of integrating text and visualization in storytelling: linking and layout. Linking is a technique that explicitly links narrative text with its explanatory visualization element and vice versa. Layout is the spatial arrangement of text and visualization in storytelling. We find participants performed better in the slideshow layout on comprehension tasks. Linking can significantly improve the user engagement in terms of user-perceived ratings, time spent on reading story, and in-depth interaction numbers. We explore users' reading behaviors across different conditions. Our findings suggest new opportunities for integrating text and visualizations in visual storytelling.

## 9. Acknowledgments

We thank the HCI lab at University of Notre Dame for their feedback and constructive critiques. We thank Poorna Talkad Sukumar and Suwen Lin for proofreading and statistical analysis suggestions. This project was supported by the National Science Foundation under Grant No. 1755734.

## References

- [AES05] AMAR R., EAGAN J., STASKO J.: Low-level components of analytic activity in information visualization. In *Information Visualization, 2005. INFOVIS 2005. IEEE Symposium on* (2005), IEEE, pp. 111–117. 4
- [AKR15] AISCH G., KELLER J., REBECCA L.: The world according to china. *New York Times* (2015). URL: <https://www.nytimes.com/interactive/2015/07/24/business/international/the-world-according-to-china-investment-maps.html>. 1
- [AM15] AISCH G., MARSH B.: Road map. *The New York Times* (2015). URL: [https://www.nytimes.com/interactive/2015/01/29/sunday-review/road-map-home-values-street-names.html?\\_r=0](https://www.nytimes.com/interactive/2015/01/29/sunday-review/road-map-home-values-street-names.html?_r=0). 2

- [AP15] ASHKENAS J., PARK H.: The race gap in america's police departments. *The New York Times* (2015). URL: <https://www.nytimes.com/interactive/2014/09/03/us/the-race-gap-in-americas-police-departments.html>. 2
- [BC13] BARRY M., CARD B.: Visualizing mbta data. URL: <http://mbtaviz.github.io/>. 2
- [BDF15] BOY J., DETIENNE F., FEKETE J.-D.: 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* (2015), ACM, pp. 1449–1458. 1, 2, 3
- [BMG\*10] BATEMAN S., MANDRYK R. L., GUTWIN C., GENEST A., MCDINE D., BROOKS C.: 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* (2010), ACM, pp. 2573–2582. 3
- [BOH11] BOSTOCK M., OGIEVETSKY V., HEER J.: D<sup>3</sup> data-driven documents. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2301–2309. 1
- [BOZ\*14] BROWN E. T., OTTLEY A., ZHAO H., LIN Q., SOUVENIR R., ENDERT A., CHANG R.: Finding waldo: Learning about users from their interactions. *IEEE Transactions on visualization and computer graphics* 20, 12 (2014), 1663–1672. 9
- [BVB\*13] BORKIN M. A., VO A. A., BYLINSKII Z., ISOLA P., SUNKAVALLI S., OLIVA A., PFISTER H.: What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2306–2315. 3
- [CCH\*14] CARENINI G., CONATI C., HOQUE E., STEICHEN B., TOKER D., ENNS J.: Highlighting interventions and user differences: informing adaptive information visualization support. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (2014), ACM, pp. 1835–1844. 6
- [DBD17] DIMARA E., BEZERIANOS A., DRAGICEVIC P.: Narratives in crowdsourced evaluation of visualizations: A double-edged sword? In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (2017), ACM, pp. 5475–5484. 3, 4, 6
- [DSL15] DEPILLIS L., SALUJA K., LU D.: A visual guide to 75 years of major refugee crises around the world. *The Washington Post* (2015). URL: <https://www.washingtonpost.com/graphics/world/historical-migrant-crisis/>. 2
- [FBM16] FULDA J., BREHMER M., MUNZNER T.: Timelinecurator: Interactive authoring of visual timelines from unstructured text. *IEEE transactions on visualization and computer graphics* 22, 1 (2016), 300–309. 1
- [FJSA17] FESSENDEN F., JASMINE L., SERGIO P., ANJALI S.: "immigrants from banned nations: Educated, mostly citizens and found in every state. *New York Times* (2017). URL: <https://www.nytimes.com/interactive/2017/01/30/us/politics/trump-immigration-ban-demographics.html?mcubz=0>. 3
- [GP01] GERSHON N., PAGE W.: What storytelling can do for information visualization. *Communications of the ACM* 44, 8 (2001), 31–37. 2
- [HD11] HULLMAN J., DIAKOPOULOS N.: Visualization rhetoric: Framing effects in narrative visualization. *IEEE transactions on visualization and computer graphics* 17, 12 (2011), 2231–2240. 1, 2
- [HDA13] HULLMAN J., DIAKOPOULOS N., ADAR E.: Contextifier: automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), ACM, pp. 2707–2716. 1, 2
- [HJ89] HEGARTY M., JUST M. A.: 10 understanding machines from text and diagrams. In *Advances in psychology*, vol. 58. Elsevier, 1989, pp. 171–194. 2, 6
- [HR07] HEER J., ROBERTSON G.: Animated transitions in statistical data graphics. *IEEE transactions on visualization and computer graphics* 13, 6 (2007), 1240–1247. 5
- [KHA14] KONG N., HEARST M. A., AGRAWALA M.: Extracting references between text and charts via crowdsourcing. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (2014), ACM, pp. 31–40. 3, 9
- [KLK17] KONG H.-K., LIU Z., KARAHALIOS K.: Internal and external visual cue preferences for visualizations in presentations. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 515–525. 2
- [KM13] KOSARA R., MACKINLAY J.: Storytelling: The next step for visualization. *Computer* 46, 5 (2013), 44–50. 1, 2, 3, 4
- [KSJ\*14] KWON B. C., STOFFEL F., JACKLE D., LEE B., KEIM D.: Visjockey: Enriching data stories through orchestrated interactive visualization. In *Poster Compendium of the Computation+ Journalism Symposium* (2014), vol. 3. 3
- [LKS13] LEE B., KAZI R. H., SMITH G.: Sketchstory: Telling more engaging stories with data through freeform sketching. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2416–2425. 2
- [May02] MAYER R. E.: *Multimedia learning*. In *Psychology of learning and motivation*, vol. 41. Elsevier, 2002, pp. 85–139. 2
- [McC11] MCCLOUD S.: *Making comics*. Harper Collins, 2011. 9
- [MHR\*17] MCKENNA S., HENRY RICHE N., LEE B., BOY J., MEYER M.: Visual narrative flow: Exploring factors shaping data visualization story reading experiences. In *Computer Graphics Forum* (2017), vol. 36, Wiley Online Library, pp. 377–387. 1, 2, 3
- [MKK15] MAHYAR N., KIM S.-H., KWON B. C.: Towards a taxonomy for evaluating user engagement in information visualization. In *Workshop on Personal Visualization: Exploring Everyday Life* (2015), vol. 3, p. 2. 3
- [MM99] MANI I., MAYBURY M. T.: *Advances in automatic text summarization*, vol. 293. MIT Press, 1999. 9
- [MTP13] MASON L., TORNATORA M. C., PLUCHINO P.: Do fourth graders integrate text and picture in processing and learning from an illustrated science text? evidence from eye-movement patterns. *Computers & Education* 60, 1 (2013), 95–109. 2
- [MZJS18] METOYER R., ZHI Q., JANCZUK B., SCHEIRER W.: Coupling story to visualization: Using textual analysis as a bridge between data and interpretation. In *23rd International Conference on Intelligent User Interfaces* (2018), ACM, pp. 503–507. 3, 9
- [NS00] NORTH C., SHNEIDERMAN B.: Snap-together visualization: can users construct and operate coordinated visualizations? *International Journal of Human-Computer Studies* 53, 5 (2000), 715–739. 2
- [OT10] O'BRIEN H. L., TOMS E. G.: 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. 3
- [RF07] RODDEN K., FU X.: Exploring how mouse movements relate to eye movements on web search results pages. In *SIGIR Workshop on Web Information Seeking and Interaction* (2007), pp. 29–32. 9
- [SES16] SAKET B., ENDERT A., STASKO J. T.: Beyond usability and performance: A review of user experience-focused evaluations in visualization. In *BELIV* (2016), pp. 133–142. 3
- [SH10] SEGEL E., HEER J.: Narrative visualization: Telling stories with data. *IEEE transactions on visualization and computer graphics* 16, 6 (2010), 1139–1148. 1, 2, 3, 9
- [SH14] SATYANARAYAN A., HEER J.: Authoring narrative visualizations with ellipsis. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 361–370. 1, 2
- [SLRS17] STOLPER C. D., LEE B., RICHE N. H., STASKO J.: Emerging and recurring data-driven storytelling techniques: Analysis of a curated collection of recent stories. Retrieved in January (2017). 2

- [SMWH17] SATYANARAYAN A., MORITZ D., WONGSUPHASAWAT K., HEER J.: Vega-lite: A grammar of interactive graphics. *IEEE transactions on visualization and computer graphics* 23, 1 (2017), 341–350. [1](#)
- [SSB15] STUSAK S., SCHWARZ J., BUTZ A.: Evaluating the memorability of physical visualizations. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), ACM, pp. 3247–3250. [3](#)
- [TA83] TOWNSEND J. T., ASHBY F. G.: *Stochastic modeling of elementary psychological processes*. CUP Archive, 1983. [6](#)
- [Tse16] TSE A.: Why we are doing fewer interactives. *The New York Times* (2016). URL: <https://github.com/archietse/malofiej-2016/blob/master/tse-malofiej-2016-slides.pdf>. [9](#)