Learning Cues to Improve the Understanding of Explanatory Storytelling

Alark Joshi
Visualization and Graphics Lab
Department of Computer Science
University of San Francisco
San Francisco, CA
apjoshi@usfca.edu

Abstract—In this paper, we present learning cues in the form of a worksheet to educate and engage students about the various aspects of storytelling. These learning cues can be used in a data visualization course after students have a good understanding of data visualization fundamentals. Through these learning cues, we draw the attention of students to the various data, color, interaction, visualization, and technical decisions as well as constraints that storytellers face when constructing a data-driven story.

Index Terms—storytelling, engagement, pedagogy, learning, interaction

I. INTRODUCTION

Data-driven storytelling [1] has been widely used to explain complex concepts [2], [3], thought-provoking stories [4], [5] as well as entertaining topics [6], [7]. This data-driven storytelling wave [8] is further spurred by the possibilities afforded by interacting with these stories on phones and tablets [9]. Data-driven stories are roughly categorized into explanatory stories and exploratory stories. Explanatory stories provide a linear (designed-directed path) for the readers of the story, whereas exploratory stories allow viewers to “find themselves” [1] in the data through active engagement [10], [11].

When students learn about data-driven storytelling, they frequently learn by looking at examples of a variety of data-driven stories, but they may inadvertently miss out on appreciating the finer details that master storytellers spend effort on.

In this paper, we propose the use of specific learning cues in a data visualization course to draw attention to various elements of an interactive story for explanatory purposes. We specifically chose explanatory storytelling due to its narrative style/flow/direction as compared to exploratory storytelling where the viewer has various directions in which they may experience the story along with some personal findings/lessons learned through the exploration. We were inspired by the Five Design-Sheet methodology [12] that uses a “paper-based lo-fidelity” approach to help participants with ideation and planning.

The learning cues introduced here are to teach students in the field of data-driven storytelling about the various aspects that go into the creation of a story. Examining an interactive story through the lens of these learning cues could help students gain a deeper understanding of the technical aspects of the story as well as the content of the story. While the work is preliminary, there may even be some effect on their overall engagement [13] with the content of the story through the learning cues.

II. RELATED WORK

A. Interactive Storytelling

With the advent of interactive data visualization on the web, storytelling has evolved from text and static imagery/animations to interactive stories that inform and narrate the story in a variety of ways. Segel and Heer [14] analyzed interactive data-driven stories and provided a design space analysis for the various ways in which stories could be told. They grouped their analysis based on the genre, the visual narrative, and narrative structure. Claes and Vande Moere [13] examined the impact of narrative structure in user engagement for visualizations created to inform the public. They found that narrative structure plays a crucial role in drawing the viewer in and led to “deeper, more personal reflection of data.”

B. Engagement

Engagement in the content of a visual representation [15] or a data-driven story is one of the primary goals for the creators. Mahyar et al. [16] identified a five-step taxonomy that allows designers to measure varying levels of engagement from just viewing a visualization all the way to making a decision based on the visualization through interaction, analysis, and hypotheses testing. Hung and Parsons [15] introduced VisEn-gage - a self-assessment questionnaire that requests viewers to rate a visualization on a variety of characteristics such as aesthetics, captivation, creativity, interest, novelty, and so on to increase engagement. Boy et al. [17] explored the impact of narrative storytelling on exploratory visualizations. They added an initial story to inform participants before they interacted with one of three exploratory visualizations developed by their team. They measured user engagement by counting the number of click interactions, hover interactions, and time spent on the webpage. They found that participants who experienced the narrative components did not engage more with the exploratory stories as compared to their counterparts.

This paper has been peer-reviewed and accepted to VisActivities: 2nd IEEE VIS Workshop on Data Vis Activities to Facilitate Learning, Reflecting, Discussing, and Designing, held in conjunction with IEEE VIS 2021, New Orleans, LA, USA. 2021 - Workshop organizers: Samuel Huron, Benjamin Bach, Georgia Panagiotidou, Mandy Keck, Jonathan C. Roberts, Sheelagh Carpendale, http://visactivities.github.io
- who did not experience any narrative component before interactive with the stories.

C. Memorability

Memorability is a form of ‘deeper engagement’ [13] that designers and storytellers aim for when designing static/interactive visualizations. Research from Borkin et al. [18] found that color and human recognizable objects increase the memorability in charts as compared to traditional ‘analytical style’ [19]. Stusak et al. [20] found that physical visualizations that participants could touch and play with lead to better memorability (increased recall after two weeks).

Recent work by Obie et al. [21] examined the impact of “author-driven narratives” on the understanding and memorability of data-driven stories. They found that participants who were shown the story with the author-driven narratives understood the content in the story better than the group that did not see the narratives. They did not find any difference in the long term recall of the content in the visualizations for either approaches.

III. Approach

Our goal is to draw attention of students to the complex craft of explanatory storytelling. We are not aiming to increase the memorability or engagement with the content of the story, but it may be a side effect that requires further investigation.

The learning cues-based worksheets can be introduced in week 11 of a 15-week introductory course on data visualization. In our case, the course contains Computer Science and Data Science students (undergraduates and graduates). By week 11, students will have been exposed to the various aspects of data-driven storytelling such as various visualization techniques, interaction techniques, appropriate use of color, and so on. The worksheet satisfies the following two learning outcomes for the course: (i) Explore issues surrounding visual integrity for visual representations of data and, (ii) Familiarize students with the various kinds of storytelling techniques, data sources, and visualization techniques used.

Students will be asked to fill in the worksheet in class for the Bussed Out [4] story and the “How the Pandemic Got Out [3]” story. Students will then be asked to exchange filled-in worksheets with their partner and then discuss their observations about the data-driven story.

In the next few sections, we describe the details and rationale for the various components of the worksheet\(^1\). We identified the following categories of a story and identified individual questions for each category.

- Storytelling technique used
- Data - What kind of data was available and what editorial decisions did the storyteller have to make when creating this story?
- Emotional response elicited from the story
- Color - color scale, legend, colors being semantically resonant
- Interaction elements
- Visualization techniques
- Other details - annotations, animated transitions, and so on

In the next few subsections, we describe each category with the rationale (and examples) for the questions in each category.

A. Storytelling technique

We examined previous work on the analysis of data-driven storytelling conducted by Stolper et al. [22] as well as Segel and Heer [14]. Based on previous work, we identified four common storytelling techniques.

In this category, the student must examine and identify one of the four storytelling techniques used in the data-driven story:

1) **Scrollytelling** - Explanatory storytelling is frequently communicated using scrolling especially due to its compatibility with phones and tablets (scrollytelling). Viewers scroll up [23]–[25] or rightwards [26], to learn more about a specific topic as the designer conveys information to them. Today, designers have many tools and libraries at hand to incorporate scrollytelling [27].

2) **Stepper / Slideshows** - This technique is also used to walk the viewer through various stages of a story one frame at a time. The designer frequently adds complexity or presents a different perspective of the story in each frame. Each frame could be static [28]/interactive [29], [30]. Atlas of Emotions [31] contains stepping down to the next concept or navigating sideways to explore a different emotion.

3) **Data Videos** - While this technique is not as prevalent as the previous two techniques, data videos [32], [33] are used in combination with interaction to narrate the story to the viewer. An excellent example of this technique is “The Fallen of World War II [34].” Through the use of the video narrative, the designer draws an emotional response of sadness and shock at the scale of deaths suffered by various countries in the war. Other examples include a combination of videos and small multiples [5], [35] to convey the scale of the story.

4) **Interactive Maps** - This technique is frequently used for exploratory visualization where a viewer is given the choice to explore a certain region or story. For explanatory visualization, interactive maps are used as intermediary stages for examination and to convey information. An excellent example of this is the Misato Town [36] experience that uses scrolling in an innovative manner to inform the viewer about various aspects of the ‘town.’ Other examples of interactive maps for storytelling are available online [37].

B. Data

This category requires students to consider what kind of data was available and what editorial decisions did the storyteller have to make when creating this story. Specifically, we ask the viewer to answer the following three questions.

\(^1\)The worksheet is provided as supplementary material along with this manuscript.
1) **What kind of data was used?** List some data characteristics - This would include identifying quantitative variables (discrete/continuous, interval), qualitative variables (nominal/ordinal), and/or geographic characteristics (latitude longitude, names of geographic locations, etc.).

2) **What data is being shown and what data is not being shown?** Frequently, storytellers have to examine the data and decide what they are going to show/highlight from the data. Nami Sumida of the San Francisco Chronicle particularly liked this question and said “I have found that deciding which variables are most important, which are secondary and which are not needed is an important step when you have a large dataset.” This includes making difficult decisions about variables (or aspect of the data) that they do not show as well. Reflecting on this process could help storytellers when they are designing their own data-driven stories.

3) **What are the additional data you wish the author had included (if any)?** As a consumer of a story, you may find yourself wondering whether there were other aspects of the story that could be highlighted if they had more data or had shown certain other attributes in the data.

C. **Emotional response**

Storytellers are frequently aiming to do more than just inform the reader of the story. Master storytellers frequently evoke a strong emotional response from the viewer [38]. In this category, we ask students to reflect and report on the emotional response that the story elicits in them as they interact with the story. Students can pick from one of the following emotional responses:

- Empathy [4]
- Distaste, Disgust, Frustration [34], [39]
- Call to Action [5]
- Information / Enlightening the audience [3], [24], [26], [29]
- Questioning beliefs and behavior [5], [25]
- Increased engagement through users finding themselves in the data [10], [11]

D. **Color**

In this category, we asked students to examine the use of color in the story. Students are asked to notice the color scales used in the various visual representations and critically evaluate whether the scale used is appropriate for the data. Here are the questions in this category.

- Is the color legend clearly visible and legible?
- Do they use appropriate color scales (Colorbrewer [40], Coolororical [41] for the data being shown? Do they avoid the use of the rainbow color map [42]?
- If the appropriate color scale is not being used, comment on what color scale would be appropriate.
- Are the colors semantically resonant [43]?

E. **Interaction elements**

Even though the learning cues are focused on explanatory storytelling, interaction plays a huge part in experiencing a data-driven story. In this category, we ask students to identify all the different types of interactions used in the story (on desktop/mobile). The list contains a combination of standard interaction elements such as hover, scroll, and so on along with the interaction paradigms suggested by Heer and Shneiderman [44]. In the table shown below, we show all the different interaction elements a viewer can select when examining a story.

<table>
<thead>
<tr>
<th>Hover</th>
<th>Select (Button/Tap)</th>
<th>Scroll</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pinch and Zoom</td>
<td>Timeline Slider</td>
<td>Brush</td>
</tr>
<tr>
<td>Filter</td>
<td>Sort (ascending/descending)</td>
<td>Drag and Drop</td>
</tr>
<tr>
<td>Details-on-demand</td>
<td>Coordinated/Linked Views</td>
<td>Other</td>
</tr>
</tbody>
</table>

F. **Visualization techniques**

Visual representations of data form a crucial part of data-driven storytelling. Students familiar with the field of data visualization must be able to examine a story and identify the various visualization techniques used by the storyteller to convey their message. Students must observe the choices made to convey the data as it relates to the intended audiences’ visualization literacy [45]. Students are asked to identify the category and specific visualization technique by referring to the Financial Times Visual Vocabulary [46]. The categories in the FT visual vocabulary are as follows:

- **Deviation** - diverging bar, diverging stacked bar, spine, and surplus/deficit filled line
- **Correlation** - scatterplot, column + line timeline, connected scatterplot, bubble plot, heatmap
- **Ranking** - ordered bar, ordered column, ordered proportional symbol, dot strip plot, slope, lollipop, bump chart
- **Distribution** - histogram, dot plot, dot strip plot, barcode plot, boxplot, violin plot, population pyramid, cumulative curve, frequency of polygons, beeswarm
- **Change over time** - line, column, column + line timeline, slope, area chart, candlestick, fan chart (projections), connected scatterplot, calendar heatmap, priestley timeline, circle timeline, vertical timeline, streamgraph
- **Magnitude** - column, bar, paired column, paired bar, Marimekko, proportional symbol, isotype (pictograms), lollipop, radar, parallel coordinates, bullet, grouped symbol
- **Part-to-whole** - stacked column/bar, Marimekko, pie, donut, treemap, voronoi, arc, gridplot, venn, waterfall
- **Spatial** - basic choropleth, proportional symbol, contour map, equalized cartogram, scaled cartogram, dot density, heat map
- **Flow** - Sankey, waterfall, chord, network

G. **Other details**

There are many strategies that storytellers used to draw your attention to and keep you engaged with the content of a story. Students are asked to notice these details such as:
Annotations - Annotations are used to help Students read a story and understand the various layers associated with it. Stories such as the “What’s Really Warming the World [25]” provide excellent annotations to the viewer as they scroll through the story. Similarly, in the connected scatterplot used in the New York Times’ story “Driving Safety, in Fits and Starts [47],” they annotate the chart to teach viewers how to read the connected scatterplot.

Animated transitions - These are used widely to preserve context for viewers and add complexity to the story. Viewers are led through the story step-by-step by using transitions. The New York Times story “How the virus got out [3]” uses animated transitions to inform the audience of the spread of the Coronavirus that led to the COVID-19 pandemic.

Informative Audio/Video - Storytellers occasionally use audio/video to inform viewers (and increase audience engagement with the content at times). “Bussed Out [4]” by the Guardian is an excellent example where they show videos to tell you the personal account of homeless individuals who were bussed out to different cities in the United States. It helps the viewer see a different perspective of the complex issue surrounding bussing homeless individuals.

Sonification - Sonification [48] is a rare but unique strategy used by storytellers to use the audio channel of viewers to augment their experience. Tulp [49] used sonification to convey the rate at which the pandemic has been spreading around the world.

Informative Text (Introduction, Methodology, Summary) - There are textual aspects to storytelling that draw in a viewer and help them understand the overall context of the story. The introduction and/or summary are important, but the methodology section is key to convey information regarding data gathering, cleaning, tools used, challenges faced, experimental prototypes, and so on. A recent book by Bremer and Wu [50] provides an excellent behind the scenes look at how they worked on data-driven stories along with various hand-drawn sketches, screenshots of early prototypes, and final results.

H. Critique

In addition to observing various aspects of a story, reflecting on other aspects of the story through a critical lens can further increase the understanding of the story and its content. This category is different from the rest and requires the students to possess a higher level of competence to be able to understand and critically evaluate the story.

These are the questions that students have to fill in as they experience the story.

• Keeping in mind the challenges associated with creating a story, identify some limitations/constraints that may have existed during the creation of the story. - This question requires a viewer to be empathetic towards the constraints (data, technology, etc.) that the storyteller may have faced when developing the story.

• How could this story be improved? What in the story may be misleading? If you have any questions that are unanswered at the end, what are they? Comment on the pacing and the continuity of the narrative. This requires students to have a keen eye on strategies that may mislead the viewer as they were reading the story. These strategies may be inadvertently used (or not) and it is important for students to be aware of them. Krist Wongsuphasawat recommended asking students about the pacing and continuity of the story.

• Do the authors use the visualization techniques correctly? Do they violate any known visualization best practices? If the design decisions made in the story seem wrong, how would you do it differently? This question specifically focuses on the visual representations used, and the challenges associated with using the wrong techniques to represent data. As Cairo [51] has pointed out, there are many ways in which a chart can be used unethically to lie to the consumer of a chart.

• What other interaction techniques could be used to improve the user experience? The viewer may think of interaction techniques that may be better served to convey information or engage the audience.

• (If there is a mobile version) What are the differences between the mobile and the desktop version of the story? Increasingly, stories are being developed for consumption on desktops as well as mobile platforms [9]. If a story can be consumed on a mobile platform, this prompt asks students to state the challenges they perceive with adapting the storytelling experience to that platform. Krist Wongsuphasawat from AirBnB suggested this question based on his experience designing data visualizations in industry.

• What questions do you have for the creator of the story? This question is designed to allow students to think beyond just the information being presented to them. Can they identify different perspectives of the story that were not considered? Are there other tools/libraries that they would recommend to the storyteller? Did the designer mean to evoke certain emotional responses or were they unintentional? Such questions and many more could increase the viewer’s understanding (and possibly
even their engagement) of the story.

IV. SUMMARY

We have presented learning cues that can be used to draw attention to the multitude of decisions that storyteller makes when designing a data-driven story. These learning cues include a combination of the storytelling technique used, considerations associated with the kind of data being used, the emotional response elicited from the viewer, decisions made related to appropriate use of color, visual representations, interaction techniques used, as well as details that are unique to certain stories (animated transitions, annotations, and so on). The Critique section requires students to further examine other aspects of the story and reflect on things they may have done differently if they were tasked with designing the data-driven story. While engagement in the story may be a side effect of the learning cues, the goal is to help students appreciate the complexity involved with creating intricate, informative data-driven stories.

ACKNOWLEDGMENTS

The author would like to thank Shirley Wu (Independent Data Visualization Creator and Data Sketches, etc.), Nami Sumida (San Francisco Chronicle), and Krist Wongsuphasawat (AirBnB) for their feedback on the various drafts of the worksheet.

REFERENCES


