The Effect of Information Visualization Delivery on Narrative Construction and Development

Donia Badawood[1] and Jo Wood[1]

[1]giCentre, City University London, United Kingdom

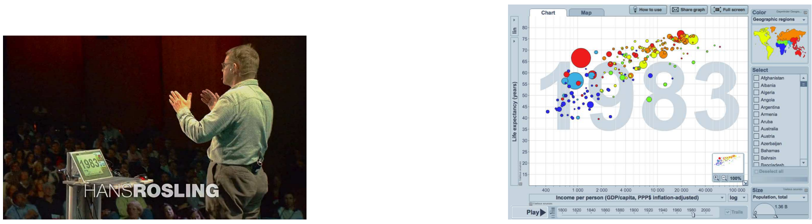

Figure 1: *From left to right: Hans Rosling representing facts on global health and wealth (narrative model), and the Gapminder World interactive information visualization software (software model)*

Abstract

We conducted a between-subject experiment with 32 participants to explore how two different models of information visualization delivery influence narratives constructed by audiences. The first model involves direct narrative by a speaker using visualization software to tell a data story, while the second model involves constructing a story by interactively exploring the visualization software. We used an open-ended questionnaire in a controlled laboratory settings in which the primary goal was to collect a number of written data stories derived from the two models. The participants' data stories and answers were all analysed and coded using a number of themes, including insight types, and narrative structures. Our findings show that while the delivery model does not significantly affect how easy or difficult the participants found telling a data story to be, it does have an effect on the tendency to identify and use outliers insights in the data story if they are not distracted from this by direct narration, and on the narrative structure and depth of the data story. Our approach to data analysis and different storytelling axes can be usefully applied to other studies and comparisons of storytelling approaches.

Categories and Subject Descriptors (according to ACM CCS): H.5.m [Information Interfaces and Presentation]: Miscellaneous—

1 Introduction

Storytelling has been a common method of communication for a long time. Many scholars have demonstrated the power of storytelling as a means of information transfer [GP01] [TKB06]. Storytelling has been extended from its origin as an oral tradition into new fields/forms of communication such as novels, films and computer games. Each of these has its own storytelling strategies, theories and techniques. Recently, a great interest in storytelling through information visualization has arisen, including two workshops on telling stories with data at VisWeek conferences in 2010 and 2011. A number of papers on the topic have also been published [HD11] [SH10] [KM13] [HDR*13]. In this study, we conducted a between-subject experiment using two different models of information visualization delivery to empirically examine how non-expert general users understand, construct and tell data stories. The first model involves watching a video in which a speaker tells a data story using information visualization software, while the second model lets users interactively explore the data using visualization software.
Although good studies have been conducted on storytelling

in information visualization, the majority of these studies are designed as case studies or theoretical arguments [HD11] [SH10] [KM13]. The main contribution of this paper is that it takes the work done in this area a step further by providing empirical results on the effect of the manner in which stories are delivered. Specifically, the paper focuses on the effect of stories delivered using information visualization by a speaker and those constructed by users as a result of an interactive exploration of information visualization. The study poses several questions about the audience's and users' constructed, written narratives such as:

- What types of insights do they gain/select to tell a story?
- How do they structure their constructed stories?
- How easy or difficult did they find telling a data story to be after experimenting each delivery model?

2 The Experiment

2.1 Experimental Factors and Questionnaire Tasks

The aim of this study is to explore and compare the effects of two different information visualization delivery models on people-constructed narratives. The InfoVis software we used to explore this was Gapminder [Gap]. This animated bubble chart includes x- and y-axes that allow user-selected variables to be compared, and the bubbles represent countries. These bubbles are coloured by continent, and an animation and/or timeline slider can be used to show how the bubbles move over time. Within this context, we examined two delivery models of InfoVis:

1. Direct narratives by a speaker using information visualization software to tell a data story to the audiences (narrative model).

For this model, we chose a video by Hans Rosling using Gapminder software to give a talk [Ros08]. Rosling's mode of storytelling with information visualization is one of the most famous in the field. He used Gapminder's animated bubble chart to give talks on several topics. The video we chose concerns child mortality. The x-axis is the income-per-person in USD, and the y-axis is the child mortality rate [Ros08].

2. Let the users/audiences explore the data interactively using the visualization software to construct data stories (software model).

For this model, we used Gapminder World software [Gap] and let the participants explore the same dataset on child mortality.

Participants were required to answer five questions after watching the video (narrative model) and five equivalent questions after exploring the data on Gapminder (software model). As we are not measuring the usability of the delivery models, measures such as time and accuracy would not be appropriate; instead, we are interested in how each model affects users in constructing narratives and telling data stories. Hence, open-ended questions are important. Moreover, open-ended questions in which participants tell data stories in writing help them formulate their mental models about the story and produce unitary narratives [NSD11]. This is in contrast to think-aloud techniques used in insight-based evaluation, which generates a series of insights gained in the order they were discovered [Nor06] [NSD11]. The main question, Q2, was asked to trigger the re-telling of a story. Re-telling is a widely used task in education to assess comprehension [FNHM09]. The five open-ended questions on each delivery model are shown in Table 1.

2.2 Participants and Experimental Procedure

A between-subject experimental design was used. Three selection criteria were identified: participants who had not taken a data visualization course, who did not have advanced knowledge in information visualization, and who were not professional data analysts. In other words, we aimed for educated but non-expert information visualization users. We recruited 32 subjects (13 females and 19 males) ranging from 22 to 56 years old from a local university. Subjects were assigned to two groups, with one group for each delivery model. Each group had 16 subjects balanced by gender. The entire experiment was carried out in a single session for each participant. The total participation time for a single participant was about half an hour. Initially, each participant was briefed on the purpose of the experiment and the experimental procedure and was asked to sign a consent form. Group I watched a 10-minute video of Hans Rosling presenting data on child mortality using an animated bubble chart [Ros08]. Then, they answered the five questions on the narrative model.

Group II was briefed about Gapminder and interactively explored a dataset on child mortality (software model), which was the same as the dataset used in the video in the first delivery model. Participants were asked not to change the indicators (x- and y- axes) when exploring the data in order to control the number of indicators the participants had to work with in both groups. Then, they answered the five questions on the software model.

After the experiment, for each delivery model, the participants were asked to answer two five-point Likert-scale questions. The two questions on Hans Rosling's video were as follows:

1. How easy or difficult did you find telling a story after watching the video?

2. How curious were you about the data/story in the video?

The answeres ranged from "*easy*" to "*difficult*" for the first question and from "*not at all*" to "*very curious*" for the second. Two equivalent questions were asked about the data/story explored in Gapminder.

3 Data Analysis

In this section, we describe the process of qualitatively analysing and coding the data; in the next section, we report the generated hypotheses and the findings of the study. Due to limited space, we only detailed the qualitative analysis and

Table 1: *Questions used for each delivery model in the experiment*

Narrative Model	Software Model
1. What was the video mostly about? (Approx. 1-2 min)	1. What was the data you explored in Gapminder mostly about? (Approx. 1-2 min)
2. Re-tell the story you gained from the video in as much detail as you can. Try to write a story that makes sense to someone who is not familiar with the story/topic. (Approx. 6-8 min)	2. Re-tell the story you gained from Gapminder in as much detail as you can. Try to write a story that makes sense to someone who is not familiar with the story/topic. (Approx. 6-8 min)
3. What did you learn that you did not already know? In other words, describe new information/knowledge you gained from the video. (Approx. 2-3)	3. What did you learn that you did not already know? In other words, describe the new information/knowledge you gained from Gapminder? (Approx. 2-3 min)
4. Did you learn something that contradicts what you already know about the topic? What is it? (Approx. 2-3 min)	4. Did you learn something that contradicts what you already know about the topic? What is it? (Approx. 2-3 min)
5. What do you think the speaker's purpose was in producing this video? (Approx. 2-3)	5. What do you think the purpose was in providing this data in Gapminder? (Approx. 2-3 min)

findings from answers to the main question used in the experiment: Q2 (re-telling the data story). Qualitative coding was processed as follows. We coded the data iteratively until generating all themes/codes. After maintaining all themes and categories and coding the data, a codebook was created. Q2 (re-telling the story) was the question expected to generate the most insights, and the answers to that question were the longest. Therefore, we randomly chose a sample of nine participants' stories and distributed them, along with the codebook, to two colleagues from the same research centre to code separately. The percentage of agreement between all coders was generally good, ranging from 81.48%-100%. Although the inter-coder agreement was good, we looked at the possible sources of disagreement and refined the codebook by adding more examples and clarifications to avoid any confusion. This refinement was discussed between coders, and they agreed on the codes generated based on that refinement. The revised codebook was then used to code the rest of the data.

3.1 Q2: Re-Telling the Data Story

In our analysis of the answers to this question, participants' stories were coded based on two themes: *Insight Type* and *Narrative Structure*.

3.1.1 Insight Type Theme

When analysing the participants' stories based on this theme, we were inspired by insight-based evaluation [Nor06] [NSD11] [Nor05]. However, we only identified the insight types reported by each participant but did not quantify their number, as it is difficult to count insight occurrences in written stories where a sentence may contain more than one insight type.
The insight types that emerged and were used in coding the data based on the *Insight Type* theme were as follows.

- *General Pattern:* the general trend or pattern of most countries.
- *Outlier:* maximum, minimum or anything outside of the general pattern; in other words, an exception.
- *Trade-Off:* a combination of minimum and maximum or making comparisons between most and least in terms of one or more specific factors.
- *Grouping:* to group different things in one category based on one or more specific criteria. In other words, define a subset or category of data.
- *Detailed Pattern:* description of details on specific points of time or instances in general patterns.

3.1.2 Narrative Structure Theme

The second theme used to analyse Q2 answers was the narrative structure. The structure is simply how the story progressed. In our analysis, we recorded up to two structures for each story. All stories were assigned *a main narrative structure* that was the most appropriate to explain the narrative flow in the story. Some also had *a sub-structure* that was the structure used within the main story theme. For example, the main structure might have been a chronological one; but within this chronology, it may have demonstrated a cause-and-effect over a specific period. Some of the stories written by the participants had a clear sub-structure, while others had only a main structure. Each of these main and sub-structures was assigned one of the following five types.

- *Problem-Solution:* emphasizing the problem and suggesting solutions whether from external information used by the narrator in the first delivery model or from personal knowledge about the topic.
- *General-to-Specific:* starting from the general trend to more specific instances, details, outliers, and more specific insights and relationships.
- *Specific-to-General:* starting from outliers, maximum, minimum, etc., to general trends or the big picture.
- *Chronological:* starting from past to present, present to past or using time points/intervals to control story progression.

- *Cause-and-Effect:* describing a figure, pattern or insight and providing the cause(s) for this insight. This structure is beyond the simple correlation between two factors (x and y-axes) and involves richer explanations of causes.

4 Hypotheses Generation and Findings

In this section, we report the results of the data analysis. We first analysed the data qualitatively as described in section 3. After the qualitative analysis, we quantified the codes generated from the qualitative analysis to look at differences and patterns in more details, which is a widely used approach in qualitative data analysis. Then, we tested a number of hypotheses as shown in sub-sections 4.1 to 4.3 using Fisher's exact test (an alternative to χ^2 for small samples).

4.1 Insight Types Reported/Used in Data Stories

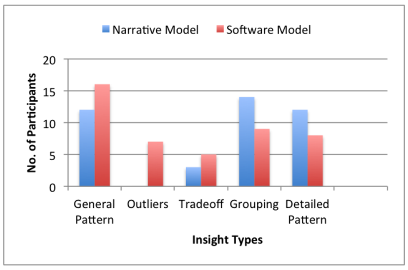

Figure 2: *The number of participants who reported each insight type in their stories for each delivery model*

A summary of the insight types gained and used in the participants' stories constructed with each delivery model is shown in Figure 2. We noticed that seven participants out of the 16 who used the software model reported an outlier insight in their stories, while the whole group who used the narrative model did not report any outliers. So, we tested the following hypothesis:

H1: The InfoVis delivery model has an effect on the existence of outlier insights in the participants' stories.

We found that the difference between the two delivery models in terms of the use of outlier insights in the participants' stories is very statistically significant (Fisher's exact test two-tailed p-value=0.0068).

4.2 Narrative Structure: Sequencing Story Events

Generally, the most common narrative structure in the participants' stories in both delivery models was the General-to-specific structure. Furthermore, it can be argued that both the Problem-Solution and Cause-and-Effect structures provide more depth to the data story than other structures. They involve going beyond simple correlation between x and y-axes to either explain causes for specific patterns or discuss the problem and suggest solutions. So, after qualitatively coding the data, the following hypothesis was generated:

H2: There is an association between the InfoVis delivery model and the use of a Problem-Solution and/or Cause-and-Effect narrative structures (either as main or sub-structures).

We found that 13 out of the 16 participants who used the narrative model used the Problem-Solution and/or Cause-and-Effect narrative structures in their stories, while only three out of the 16 who used the software model used these structures in their stories. The association between the delivery model and the existence of these two structures is considered to be very statistically significant (Fisher's exact test two-tailed p-value=0.0011).

4.3 Difficulty of Telling a Data Story

It can be thought that re-telling a data story after it is narrated by a speaker is certainly easier than constructing narratives as a result of interactive exploration of the visualization software. We tested the following hypothesis to prove that it is not that one delivery model makes it easier or more difficult to tell a data story; rather, it is how the delivery model may guide the audience's attention in a way that could affect the data story people get.

H3: The InfoVis delivery model has an impact on the level of difficulty of telling a data story.

A statistical test showed that the level of difficulty of telling a data story was not significantly affected by which delivery model the participants used (Fisher's exact test two-tailed p-value=0.48).

5 Conclusion and Discussion

Outlier insight was used more often by the participants who constructed data stories by exploring the data in Gapminder. The fact that outliers caught the audiences' attention is important, particularly in the absence of a narrator who directs the audience's attention to the desired events. The danger in this approach is that outliers, by definition, are not representative of the data as a whole, so care is needed to lead users from attention-grabbing outliers to the core message implied by the data.

As it is more likely that participants will use a Problem-Solution or Cause-and-Effect narrative structure with the narrative model, one should pay special attention to these kind of details/depth in the absence of a narrator who can justify, explain, or provide background information. Annotations and choice of labels play an important role in this case. Moreover, it is important to note that using a specific narrative structure by a participant (e.g., General-to-Specific) does not necessarily mean that the general insights in the beginning of a participant's story are more important than the following insights. Similarly, reporting specific insight types does not necessarily mean that these are the only insight types the participant could gain. Storytelling is selective by nature, and the reported insights represent those that contributed to the overall mental model of the audience rather than a quantification of what they gained.

Acknowledgement

We would like to thank the participants of the study, Jason Dykes for useful feedback on an earlier version of the paper, and Aidan Slingsby and Sarah Goodwin from the giCentre for qualitatively coding a sample from the data in order to help us assessing inter-coder reliability.

References

[FNHM09] FULKERSON D., NICHOLS P., HAYNIE K., MISLEVY R.: *Narrative Structures in the Development of Scenario-Based Science Assessments (Large-Scale Assessment Technical Report 3)*. Tech. rep., SRI International, Menlo Park, CA, 2009. URL: http://ecd.sri.com/publications.html#. 2

[Gap] Gapminder. http://www.gapminder.org. Accessed: 15-04-2014. 2

[GP01] GERSHON N., PAGE W.: What storytelling can do for information visualization. *Commun. ACM 44*, 8 (Aug. 2001), 31–37. URL: http://doi.acm.org/10.1145/381641.381653, doi:10.1145/381641.381653. 1

[HD11] HULLMAN J., DIAKOPOULOS N.: Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics 17*, 12 (Dec. 2011), 2231–2240. doi:10.1109/TVCG.2011.255. 1, 2

[HDR*13] HULLMAN J., DRUCKER S., RICHE N., LEE B., FISHER D., ADAR E.: A deeper understanding of sequence in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2406–2415. doi:10.1109/TVCG.2013.119. 1

[KM13] KOSARA R., MACKINLAY J.: Storytelling: The next step for visualization. *Computer 46*, 5 (2013), 44–50. doi:http://doi.ieeecomputersociety.org/10.1109/MC.2013.36. 1, 2

[Nor05] NORTH C.: Information visualization. In *Handbook of Human Factors and Ergonomics*, 3rd ed. John Wiley & Sons, New York, 2005. 3

[Nor06] NORTH C.: Toward measuring visualization insight. *IEEE Computer Graphics and Applications 26*, 3 (June 2006), 6–9. doi:10.1109/MCG.2006.70. 2, 3

[NSD11] NORTH C., SARAIYA P., DUCA K.: A comparison of benchmark task and insight evaluation methods for information visualization. *Information Visualization 10*, 3 (July 2011), 162–181. URL: http://ivi.sagepub.com/content/10/3/162.abstract, doi:10.1177/1473871611415989. 2, 3

[Ros08] ROSLING H.: Reducing child mortality. ttp://www.gapminder.org/videos/gapmindervideos/gapcast-11-reducing-child-mortality/, 2008. Accessed: 15-04-2014. 2

[SH10] SEGEL E., HEER J.: Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (Dec. 2010), 1139–1148. doi:10.1109/TVCG.2010.179. 1, 2

[TKB06] TERGAN S.-O., KELLER T., BURKHARD R. A.: Integrating knowledge and information: digital concept maps as a bridging technology. *Information Visualization 5*, 3 (2006), 167–174. doi:10.1057/palgrave.ivs.9500132. 1