The Effect of Information Visualisation Delivery on Narrative Construction and Development

Donia Badawood* Jo Wood**
City University London, UK City University London, UK

ABSTRACT

We conducted a within-subject experiment involving 13 participants that empirically explore how two different models of story delivery involving information visualization influence audience-constructed narratives. The first model involves a speaker using visualization software to communicate a direct narrative, while the second involves constructing a story by interactively exploring visualization software. We used an open-ended questionnaire in controlled laboratory settings, with the primary goal of collecting a number of stories derived from the two models, followed by two Likert-scale questions on the ease of telling and curiosity about the story in each delivery model. We qualitatively analysed the stories constructed by the participants, based on a number of themes tied to storytelling, including time and place and narrative structure. The study's results reveal some interesting possible differences in how users receive, interpret, and create stories in each case.

Keywords: Evaluation, Narrative Visualization, Storytelling.

1 INTRODUCTION

Storytelling has long been a common way of communicating. Recently, interest in storytelling through information visualization has increased, with two workshops on telling stories with data held at the VisWeek conferences in 2010 and 2011, and a number of papers published on the topic [2, 4]. In this study, we conducted a controlled experiment to empirically examine how non-expert general users understand, construct, and tell stories, using two different models of story delivery involving information visualization. The first involved watching a video in which a presenter told a story using an information visualization design, while the second let users explore an interactive visualization program by themselves.

A literature review revealed that while good studies had been conducted on storytelling in information visualization, but the majority of the research involved case studies and theoretical frameworks [2, 4]. The lack of empirical evaluation of the effect of authored narration on the interpretation of data-rich visualization motivated the work carried out in this study.

2 THE EXPERIMENT

The aim of this study is to explore and compare the effects of two different story delivery models on individuals' methods of constructing narratives, comprehending data, and telling stories about it. The information visualization software used to explore this was Gapminder [1], which is an animated bubble chart in which users can choose selected variables to be compared as x and y axes, with bubbles representing 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 storytelling:

Model #1: Direct narratives by a speaker using information visualization to deliver a story to an audience.

For this model, we chose a video of Hans Rosling talking in TED, using Gapminder's animated bubble chart [3]. The video we chose concerns the HIV epidemic, and used an x-axis to represent income per person in USD, and a y-axis to indicate the percentage of adults infected by HIV.

Model#2 Audience-explored data using an interactive visualization software similar to that represented by the speaker in the first model.

For this model, we used Gapminder World software [1] and let the participants interactively explore a dataset on child mortality (y-axis) and fertility rate (x-axis). We chose a different dataset for each delivery model to avoid participants' answers being influenced by the first model that they had experienced.

Thirteen participants (9 females and 4 males) aged between 23 and 48 took part in the experiment. Three selection criteria were applied: participants should not have taken a data visualization course in the past, should not have advanced knowledge of information visualization, and should not be professional data analysts. The entire experiment was carried out in a single one-hour session for each participant. The participants were randomly divided into two groups. Each group was shown the two models of story delivery in a different order, so as to account for learning effect. The first group watched the video and then answered five open-ended questions:

1. What was the video 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.)
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 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 min.)

Then, participants explored data interactively on Gapminder. As the duration of the video in the first model was roughly 10 minutes, we allowed the participants to explore the data in the second model for the same length of time. We also asked them not to change the indicators (x and y axes), in order to control the number of indicators involved in both models. Then, they answered five questions similar to those above, with the only difference being the word 'video' to refer to the second delivery model of Gapminder software, rather than the first.

At the end of the experiment, participants were asked to answer two five-point Likert-scale questions on each model. The two questions on 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 answers ranged from "very easy" to "very difficult" for the first question and from "not at all" to "very curious" for the

* Donia.Badawood.1@city.ac.uk
** jwo@soi.city.ac.uk

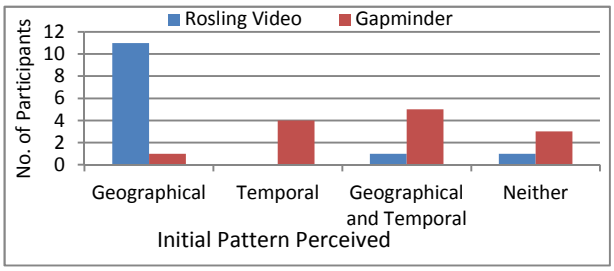

Fig.1 Initial perception of Geographical and Temporal patterns in the story in the two delivery models.

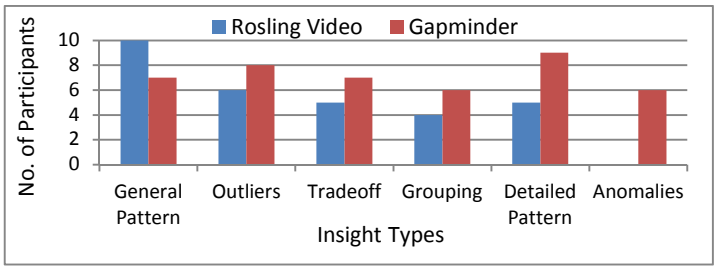

Fig. 2 The number of participants (out of 13) who reported each insight type in their stories for each story delivery model

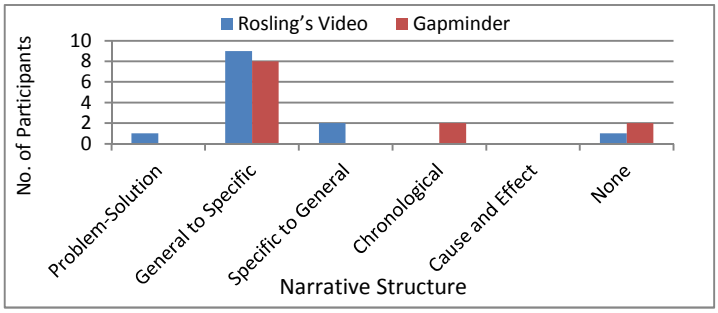

Fig 3 The number of participants (out of 13) who used each narrative structure in their stories for each story delivery model.

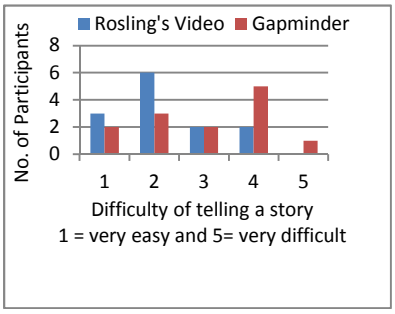

Fig. 4 Participants' subjective feedback on the level of difficulty of telling a story for each story delivery model.

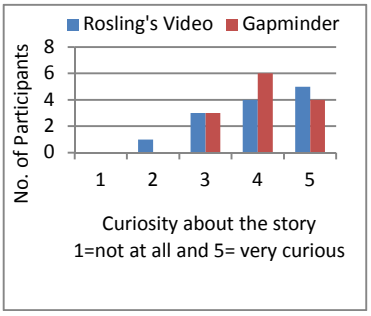

Fig 5 Participants' subjective feedback on the level of curiosity about the story in each story delivery model

second. Two similar questions were asked about the data/story explored through Gapminder.

3 Data Analysis, Results, and Discussion

A thematic analysis was used to analyse responses, using a combination of data-driven and preset themes. The answers to Q1 were coded into four categories, depending on the initial pattern(s) perceived. These categories were: *geographical*, *temporal*, *geographical and temporal*, and *neither*.

The answers to Q2 were coded based on two main themes: insight types and narrative structure (sequencing story events). The insight types used in coding data based on the insight type theme were: *general pattern*, *detailed pattern*, *outliers*, *trade-off* (a combination of maximum and minimum), *grouping*, and *anomalies* (data errors). Five narrative structures were also emerged, and used in coding the participants' stories based on the second theme, narrative structure (Fig. 3).

After watching Rosling's video, most of the participants emphasised only geographic patterns, while most emphasised both geographic and temporal patterns after exploring the data on Gapminder (Fig. 1).

The participants' stories after exploring the data on Gapminder contained a slightly greater variety of insight types than the stories after watching Rosling's video. However, an interesting exception was that the *anomalies* insight was only gained through the use of Gapminder (Fig.2), perhaps due to the fact that participants' attention was directed by Rosling in the video, making them less likely to spot deficiencies in data that were not pointed out. With regard to narrative structure, the most common structure in the participants' stories was *general to specific* (Fig.3). Two participants used the *chronological* structure with the Gapminder model. This might suggest that when participants explore time-series data independently, the role of time in structuring and initiating progression in story events becomes apparent. With regard to the expected purpose in delivering stories in each delivery model (Q5), most answers concerned raising awareness about the topic or providing data in a more compelling way. Two participants thought that there was a political or personal purpose behind the story in Rosling's video. The apparent objectivity offered by the Gapminder software did not prompt such responses. Finally, the subjective feedback questions showed that most participants found telling a story easier after watching Rosling's video. It was not clear, however, that the more curious the participants were about data or stories, the easier they found telling a story, although this was found to be true in some cases.

4 Conclusion, Limitations, and Future Work

Exploring the data interactively through the Gapminder software gave participants more insight into detailed patterns, grouping, and anomalies than did viewing Rosling's video. Our findings also suggested that participants are more likely to neutrally accept a narrative's message if they construct and make conclusions about the narrative by themselves. The main limitation of this study was the small sample size involved. Increasing the sample size would enable us to carry out more statistical analysis after qualitatively coding data. We acknowledge that there may have been confounding factors that influenced the results in some cases. Mainly that each delivery model uses a different dataset. This will be taken into account in the future work. A further complementary user study is planned, which will differ from the present experiment in that the first delivery model will be a video or presentation of the visualizations used by Rosling without him present. The findings of the two sets of experiments will then be compared.

References

[1] Gapminder. http://www.gapminder.org.

[2] J. Hullman and N. Diakopoulos, "Visualization Rhetoric: Framing Effects in Narrative Visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, no. 12, pp. 2231–2240, Dec. 2011.

[3] H. Rosling, TED 2009, http://www.gapminder.org/videos/ted-talk-2009-hans-rosling-hiv-facts/

[4] E. Segel and J. Heer, "Narrative Visualization: Telling Stories with Data," IEEE Transactions on Visualization and Computer Graphics, vol.16, no. 6, pp. 1139–1148, Dec. 2010.