Presenting Data Chart Sequentially v.s. Un-sequentially

Nina(Zhuxiaona) Wei
DUB Group, University of Washington
Seattle, WA
zwei7@uw.edu
(206)-295-0500

Christine(Yenting) Liu
Information School, University of Washington
Seattle, WA
ytliu@uw.edu
(206)-747-2905

ABSTRACT
The research focuses on the effectiveness of communicating a simple chart to audiences. The original hypothesis is that presenting a single statistic data chart sequentially works better than presenting it un-sequentially, in terms of the degree of understanding for audience. Base on the hypothesis, nine simple charts contain bar, line and scatter plot charts are chosen from *The Economist* with designed follow-up questions to conduct experiments for quantifying participants' degree of understanding toward the chart on Amazon Mechanical Turk. After analyzing 180 responses, the results turns out that presenting a chart sequentially comes out with better results. However, statistically speaking, there is no significant difference between presenting the chart sequentially and un-sequentially. We analyzed and discussed the reasons behind, and addressed our limitations and future work by the end of the paper.

Author Keywords
Narrative Visualization; chart; storytelling; sequence;

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Storytelling, regarded as the next step for visualization, has been showing its great potential in visualization research, especially, Kosara and Mackinlay estimated that visualization experts should focus on storytelling strategies in presentation [3].

There are a lot of work regarding in sequence in presentation, slideshows or other narrative visualization types. But, how can we utilize storytelling strategies in presenting a single data chart? How about the sequence

inside a chart? According to cognitive psychology, presenting the chart sequentially (Figure 1) uses more "bottom-up" processing, while presenting the chart un-sequentially, in a traditional way, uses more "top-down" processing, which may differs in graphic perception as well. Also, when presenting the chart sequentially, users can focus on each element (title, legend, axis, and content) one by one, instead of being overwhelmed by lots of information at the same time. Thus, we are motivated to explore the sequence in presenting a single data chart.
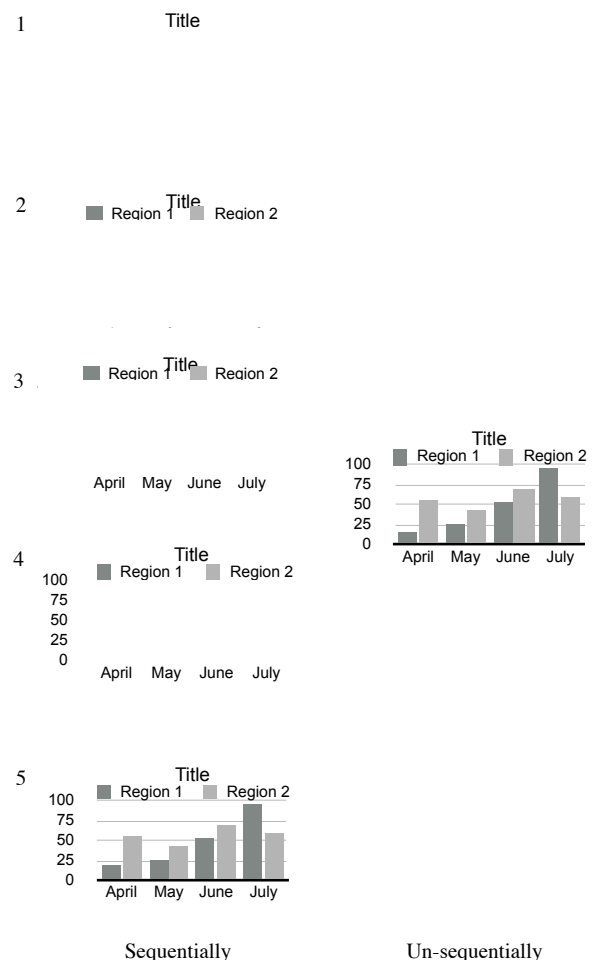

Figure 1. A chart presented sequentially and un-sequentially.

RELATED WORK

Segel and Heer analyzed storytelling strategies used in 58 collected online narrative visualization samples [4]. Followed their works, Hullman et al. specifically studied sequence, a primary element in storytelling, in slideshow-style presentation: Their study suggested that it is a need for more sophisticated global constraints than simply summing local transition costs to determine the best path through a graph of weighted visualization transitions [2].

Heer and Robertson also studies animated transitions, specifically in statistical data graphics: They extended theoretical models of data graphics and introduced a taxonomy of transition types, and also demonstrated that animated transitions can significantly improve graphical perception [1].

Above all, sequence, as well as transitions in statistic data charts has been proved can improve the effectiveness of visualization. From the above analysis, we came up with our hypothesis — presenting a single statistical data chart sequentially is better than presenting it un-sequentially. As a matter of fact, there is no related work addressing this question, so that we wanted to take the first step, to explore the narrative visualization in a single data chart. In this study, we aimed to find out whether there is any significant difference between the two presenting ways — sequentially and un-sequentially, as well as analyze other factors, such as data type (bar chart, line chart, and scatterplot).

METHODS

Pilot test

Before stepping into expeiriments, researchers first run a pilot test with 4 participants. The content of pilot study contains only bar charts. The participants are asked to read a chart both sequentially and un-sequentially, and then have a semi-structure interview to know about their understanding and interpretation. The result of the pilot study indicates that most of our participants have issue with the time has been given in the expreiment. However, considering enough time would result in similar pattern for sequentially and un-sequentially, we only extended the timing a bit. Also, participants reported that presenting the data chart sequentially do benefit their understanding in some extent. What is more, participants pointed out that they may look at other charts, such as scatterplot, in a different way, which led to different experiment design in the following experiment.

Experiment materials

We picked 9 basic data charts, including three bar charts, three line charts, and three scatterplot charts, from *The Economist*, from the year of 2013 to 2014, the present.

Experiment design

The experiment contains 2(sequentially & un-sequentially) x 3(bar chart, line chart, and scatterplot) repeated measurement design.

The experiment is conducted on Amazon Mechanical Turk. Each chart is embedded in two web pages presenting it sequentially or un-sequentially.

There are two main experiment groups:

I. Experiment group, presenting the chart sequentially.

Under sequentially condition, participants will be provided a static chart and it stays on the screen for 13 seconds.

II. Control group, presenting the chart un-sequentially.

Under un-sequentially condition, the chart is showed in the order of title (2 seconds), legends (2 seconds), x-axis (2 seconds), y-axis (2 seconds) and content of the chart (5 seconds) (Figure 2).

Figure 2. Experiment group, presenting the sample chart sequentially

After 13 seconds' reading, the chart will be hidden and questions show up. For each chart, 6 follow-up questions are provided. The first two are objective questions, one is about the axis, and the other one is about the content.

The third question asks participants for their open insights. While, the fourth question collects participants' subjective attitude towards each chart. The fifth question aim to know about participants' general view path for a single static data chart. We let them write down their opinions o towards effectiveness of a chart in the last question.

An example of the six questions are presented below:

1. What is the range of y-axis?

2. Nike's revenue was about how many times than Puma in 2008?

3. Any insight from this chart?

4. I understand this chart very well. (Strongly disagree, Disagree, Uncertain, Agree, Strongly agree)

5. What is your viewing order when you read a static chart? e.g. Start to read from x axis, y axis, legend and content. Please list them the following elements by your reading order (title=1, x-axis=2, y-axis=3, legend=4, content=5.). e.g. 13452.

6. Please write down three (adjective) words that you think are most appropriate to describe the effectiveness of a chart.
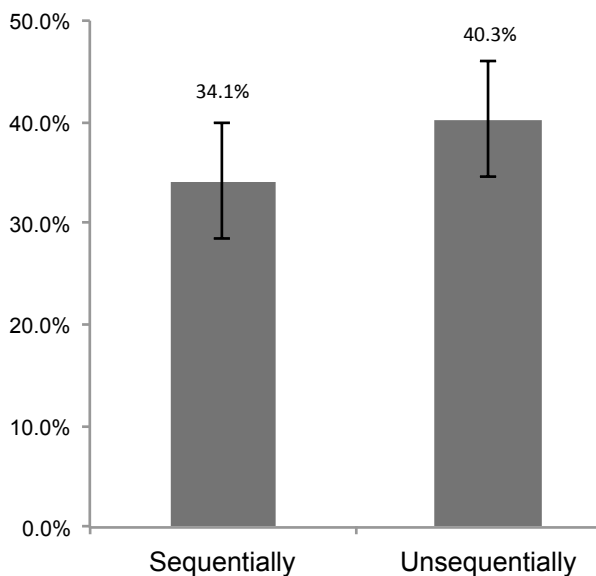

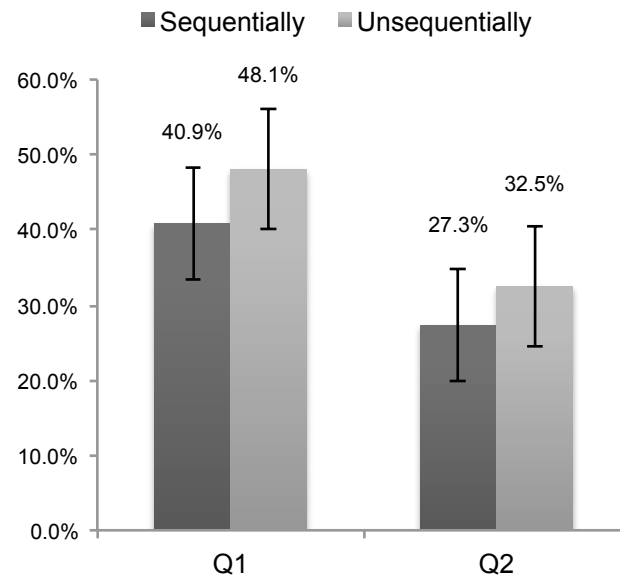**Figure 3. Accuracy of Sequentially vs Un-sequentially**

Figure 4. Accuracy of Q1 vs Q2

RESULTS

Objective interpretation

The average accuracy of Q1 & Q2, sequentially is 34.1%($\pm$0.057), and un-sequentially is 40.3%($\pm$0.059). There is no significant difference between the two ($F_{(1, 34)}=0.574$, $p=.454>.05$) (Figure 3).

Also, there is no significant difference between the accuracy of Q1, question of the axis and Q2, question of the content ($F_{(1, 34)}=3.434$, $p=.073>.05$) (Figure 4).

As for different chart type, though the average accuracy of bar chart looks higher than line chart and scatterplot, the difference is not significant as well ($F_{(2, 33)}=1.318$, $P=.281>.05$) (Figure 5).

Subjective interpretation

As for Q3, there is not many insights. Some repeated similar content of Q2, some just described the title or general information. But, when asked "I understand this chart very well", there is no significant difference between presenting the chart sequentially and un-sequentially ($F_{(1, 159)}=0.298$, $p=.586>.05$).

But, the difference among bar chart, line chart, and scatterplot is significant ($F_{(2, 158)}=5.903$, $p=.003<.01$) (Figure 6). Specifically, according to the Post Hoc Tests, the difference between bar chart and line chart is significant ($p=.015<.05$), and the difference between bar chart and scatterplot is significant ($p=.001<.05$), while, there is no significant difference between line chart and scatterplot ($p=.319>.05$).
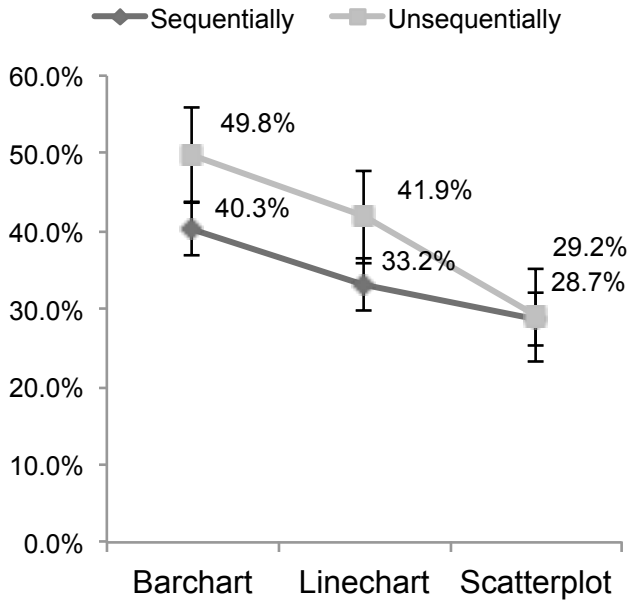
Figure 5. Accuracy of bar chart vs line chart vs scatterplot

Figure 6. Subjective scale of different charts

View path

In fact, the result of view path is extremely various. Mainly, participants follow the traditional path, from title, legend, axis to the content, but some focus on the content first.

Also, the view path differ from different chart types. For example, as for bar chart, participants prefer looking at the axis first, on the contrary, as for the scatterplot, participants tend to read the content first.

While, we noticed that a lot of the participants mistaken the question, they answered the view path of the chart we provided in the experiment, instead of a chart presenting in the real world, which may affect the result in a great extent, thus, we decided not to analyze the data of view path deeply. Though, it is true that it is much more helpful to utilize eye-tracking technique to study the view path, which we will mention in the future work.
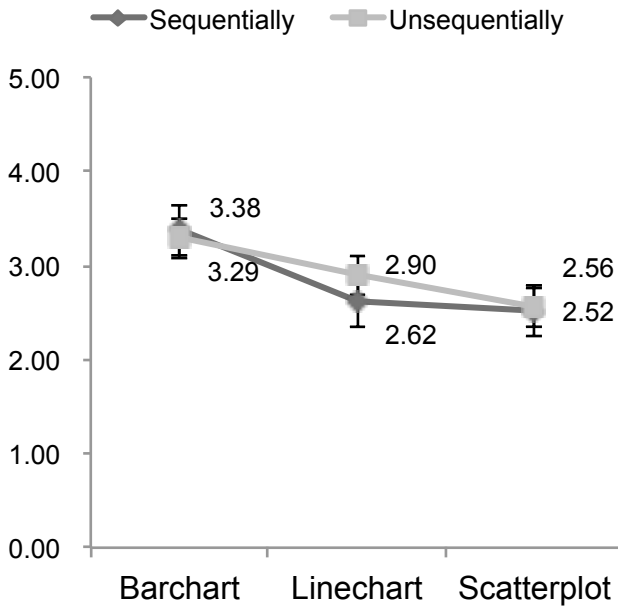

Figure 7. Word Cloud of "Effectiveness"

Effectiveness

We have assumed several words for "effectiveness", such as, engaging, interesting, attracting, but, as it turns out that general people mainly regard "clear", "simple", "concise", "easy", and "clarity" as effectiveness (Figure 7).

Though, this kind of "effectiveness" is different from the real "effectiveness" of data visualization, defined by Mackinlay, "A visualization is more effective than another visualization if the information conveyed by one visualization is more readily perceived than the information in the other visualization." [5].

DISCUSSION

In our experiment, it turns out that presenting the statistic data chart sequentially is no better than presenting it unsequentially, or even less effective, which is opposite to our hypothesis.

Though the theories — cognitive overload and perceptual processing can support our hypothesis, from the pilot test interview and follow-ups with some participants, also the communications with visualization researchers, we found out that the following points do help to explain our results.

In fact, when the user is reading a chart, probably he is not following a single and straightforward view path, but reading more than one elements at the same time. Each element in a chart is not independent but has somewhat connections with other elements. If only one element presented, it is hard for the user to catch up the whole information. It is true that some participants reported that they started to read the chart when all the elements are presented.

Another possible reason is that the charts used in our experiment are all basic and common charts — bar chart, line chart, and scatterplot, which probably do not cause

cognitive overload. Also, users are accustomed to their own traditional view path. Imaging if presenting a newly-designed or unfamiliar-style chart, such as, parallel coordinates, and voronoi diagram, maybe in a certain sequence will be better. Or, it is also will be more effective to present more and narrative interactive charts sequentially, which seems similar with utilizing animated transitions.

What is more, our initial assumption was that when the speaker is presenting his statistic data chart in a conference, it might be better to present each element one by one together with his narrative explanation. But, in our experiment, narration is not provided, which may make a difference.

All in all, though the result turns out that presenting the chart sequentially is not as effective as presenting it in a traditional way, it is not saying that adding sequence or transition into presenting the chart has totally no benefit, further study and analysis are needed.

CONCLUSION

1. Presenting a single data chart sequentially is not proved better than presenting it un-sequentially in this study.

2. There is no significant difference of objective interpretation among bar chart, line chart, and scatter plot. Bar chart is more subjectively preferred in terms of degree of understanding than line chart and scatter plot.

3. The view path varies among different type of charts. Most of participants do not read elements on charts separately. They intend to read more than one elements at a time, which reveal the useful information.

4. General people have the view that the "effectiveness" of a chart is "clear", "simple", "concise", "easy", and "clarity".

FUTURE WORK

It is only a class project within only three weeks, thus, we only walked out the first step. Our study also has limitations:

We picked the charts originally from The Economist, and did not ps too much, we though that it is the charts that people will read in their real life, which considered as the advantage, and the disadvantage is that some elements, visual or layout are uncontrollable and a bit inconsistent among different charts, which may influence the result.

What is more, the timings setting and the questions designed for evaluating the interpretation also need to be adjusted and improved.

Additionally, as the experiment was run on Amazon Mechanical Turk, we cannot know whether the worker paid full attention to the HIT without any distraction, because we only presented the chart totally for 13 seconds, which required full attention, it may leads to different results if workers were distracted.

Though, it is obvious that our study has a certain impact and there remains a lot to explore as well.

1. Use eye-tracking to study users' view path or reading pattern of a singe chart.

2. Add narration, and simulate the presentation in a real world, to study whether presenting charts sequentially has effect on the understanding and interpretation.

3. Study high-level interactive and narrative visualization, such as, Gapminder, instead of basic charts.

REFERENCES

1. Heer, J.; Robertson, G.G., "Animated Transitions in Statistical Data Graphics," Visualization and Computer Graphics, IEEE Transactions on , vol.13, no.6, pp. 1240,1247, Nov.-Dec. 2007.

2. Hullman, J.; Drucker, S.; Riche, N.H.; Bongshin Lee; Fisher, D.; Adar, E., "A Deeper Understanding of Sequence in Narrative Visualization," Visualization and Computer Graphics, IEEE Transactions on , vol.19, no. 12, pp.2406,2415, Dec. 2013

3. Kosara, R.; Mackinlay, J., "Storytelling: The Next Step for Visualization," Computer , vol.46, no.5, pp.44,50, May 2013.

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

5. Mackinlay, Jock. "Automating the design of graphical presentations of relational information." ACM Transactions on Graphics (TOG) 5.2, 1986, 110-141.