An Eye-Tracking Study on Sparklines within Textual Context

P. Ruchikachorn and P. Rattanawicha

Chulalongkorn University, Thailand

Abstract

Sparklines are placed in documents but their usability is rarely evaluated in their immediate context of paragraphs of text. We conducted an eye-tracking study to measure readability and understandability of four different conditions: two different sparkline chart types (bar and line charts) and two text languages (native and non-native languages). We found out that most participants out of 296 in total were not distracted by sparklines. Only 3.19% of the average reading time was spent looking at sparklines. There was no correlation between dwell time and data understanding, measured in a post-experiment quiz. The chart types did not have a significant effect on sparkline attention. However, compared with native textual context, sparklines in non-native text were more noticeable. The results of this study can be useful for future sparkline usage consideration.

CCS Concepts

•***Human-centered computing*** → *Empirical studies in visualization; Visualization design and evaluation methods;*

1. Introduction

Sparklines are expected to have low cognitive load due to their integration with text [BW17]. However, to the best of our knowledge, they are hardly evaluated—particularly in an eye-tracking study—to measure the interaction with their immediate context of text. Our research questions are (i) whether sparklines reduce text readability and (ii) whether chart types affect readability.

An early application of sparklines was in healthcare, especially in lieu of tabular data [PT94, PBLL12, PT13, BVM15]. They were previously evaluated by assessment time [BGB10]. Most general-purpose sparkline evaluations [GWFI14, GWBI15] studied various sparkline parameters and their effect on text legibility and readability. Because visualizations are intrinsically visual media, it is natural to evaluate them with eye-trackers [GH10, HNA*11, KFBW16].

This paper uses an eye-tracking experiment to study the effect of different textual contexts, in addition to sparklines themselves, on sparkline attention. The result can help designers or editors to consider contextual factors before adding sparklines to a document. In particular, a solution or remedy may need to be sought out if the document language is not the first language of all target readers.

2. Study Design

We sourced the text from a monthly report from the Bank of Thailand [Ban17] as the official release has both English and Thai versions. Hence, translation was not needed. The text (approximately 200 words) in full is included in the supplemental material.The presented data figures were also available in a convenient format in the Statistics page of the Bank of Thailand. The economic data in the report were the number of tourists (in millions and seasonally adjusted), inflation index (based on 2015), and export and import amounts (in billions). The selected range of the data were in 2016 and 2017. In total, there are 13 monthly data points from May 2016 to May 2017 of each data category.

We recreated the report in web standard technologies and D3.js for sparkline implementation as shown in Figure 1. (See the supplemental material for the actual size.) According to Tufte's recommendation [Tuf06], we placed a sparkline immediately after each related clause (sometimes mid-sentence), which was in the first sentence of each paragraph. The background of the first 8 data points or year 2016 is in light gray to differentiate from the last 5 data points or year 2017. Next to the first data point of each year, there is also a yearly label: 2016 and 2017 in Gregorian calendar and the equivalents of 2559 and 2560 in Buddhist calendar.

We conducted a between-subjects eye-tracking experiment whose participants voluntarily signed up for the experiment (100 male and 196 female students in a business school). Their native language is Thai but they have taken at least one college-level English course. Each participant was randomly assigned to one of four conditions and asked to read the document displayed on the computer with a Tobii Pro X2-30 eye tracker. After reading the document, besides a standard demographic questionnaire (for gender and age), there was a short quiz of three questions to test if the participant understood the data in the document. During the quiz time, the participants were not allowed to refer to or return back to the document with sparklines. The three questions targeted different levels of understanding from straightforward data reading to trend detection as follows:

DOI: 10.2312/eurp.20181119

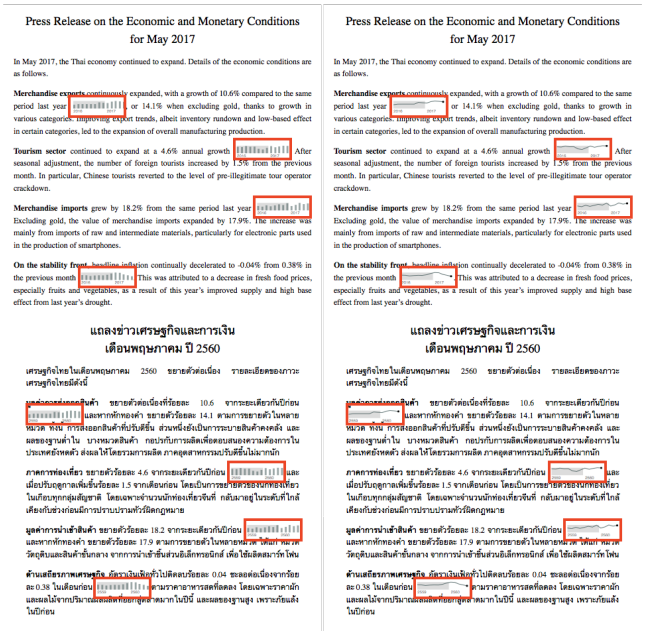

Figure 1: *Four conditions of the document in the experiment. The top and bottom rows show the document in English and Thai, respectively, while the left and right columns show the document with sparklines in bar charts and line charts. Sparklines in this figure are enclosed in red boxes to emphasize their locations; the experiment participants do not see these boxes.*

Q1: How were the import and export values during the beginning of 2017, compared with the figures of 2016?
Q2: How was the number of tourists during the end of 2016 in comparison to the beginning of 2016 and 2017?
Q3: What was the trend of inflation rate in the first half of 2017?

Each question had the same set of multiple choices of "increasing", "decreasing", "no change", and "not sure or not enough information". The correct answers to the questions were "increasing", "decreasing", and "decreasing", respectively. The first and the third sparklines could help answering the first question while the second and the fourth sparklines were useful for the second and the third questions, respectively.

Time to complete reading the document was also measured, together with each participant's eye-tracking data. These collected data were analyzed by Tobii Studio software to calculate dwell time for each area of interest (AOI). There were four AOIs defined by the rectangular areas around four sparklines in each document, shown as red rectangles in Figure 1.

Our research questions were broken down into various hypotheses. We conjectured that the participants would spend less time to complete reading text and look at sparklines in their native language because their higher reading speed (compared with reading English) may make them skip looking at sparklines. As the bar chart has more prominent positive space than the line chart, the bar chart should be more obvious to see and gain more attention. Also, the participants who pay more attention on sparklines should perform better on the quiz because sparklines should aid data understanding.

3. Results and Data Analysis

The participants spent 2 minutes 45 seconds on average to complete reading the document. For English (non-native) text, the average time to complete reading increased to almost 3 minutes, while reading Thai (native) text, the average reading time was only 2 and a half minutes. We used a one-tailed t-test to compare the mean values of time to complete reading native language (Thai) text and English text at 95% confidence interval and found out that the partcipants spent significantly less time to read Thai text ($t(270) = 3.293, p = 0.001$). However, text with different chart types had no significant differences on average time to complete reading.

The participants spent a small amount of time looking at sparklines. In other words, they were not distracted by sparklines inserted in the text they read. On average, they spent only 5.24 seconds, or 3.19% of their total reading time, on four sparklines altogether. It was less than two seconds per sparkline. A one-tailed t-test comparing the mean values of dwell time in native language (Thai) text and English text at 95% confidence interval concluded that the partcipants spent significantly less time to look at sparklines in Thai text ($t(293) = 1.909, p = 0.029$). On the other hand, we used the same t-test to show that the participants did not spend more time looking at bar-chart sparklines. In other words, the chart type does not have an effect on sparkline attention.

Also, neither sparkline chart type nor dwell time had a significant effect on the average number of correct answers, while contextual text languages significantly affected data understanding ($t(294) = 1.988, p = 0.024$). The participants who received the Thai document performed better on the quiz.

4. Conclusion and Future Work

We concluded that the participants did not get distracted by sparklines. Sparkline chart types did not affect sparkline readability but text languages did. Sparklines that were surrounded by the native language received less attention.

To answer a broader question of sparkline understandability, we will conduct another eye-tracking experiment between two conditions with and without sparklines. With a control group, we can ask the participants to provide comparative subjective ratings. As sparklines are popular and readily available, this paper hopes to help bring more consideration for not only the design of sparklines but also their surrounding context.

Acknowledgements

This work was partially supported by Grants for Development of New Faculty Staff, Ratchadaphiseksomphot Endowment Fund, Chulalongkorn University.

References

[Ban17] BANK OF THAILAND: Economic and Financial Statstics, 2017. 1

[BGB10] BAUER D. T., GUERLAIN S., BROWN P. J.: The Design and Evaluation of a Graphical Display for Laboratory Data. *Journal of the American Medical Informatics Association 17*, 4 (2010), 416–424. 1

[BVM15] BARSNES H., VAUDEL M., MARTENS L.: JSparklines: Making Tabular Proteomics Data Come Alive. *PROTEOMICS 15*, 8 (2015), 1428–1431. 1

[BW17] BECK F., WEISKOPF D.: Word-Sized Graphics for Scientific Texts. *IEEE Transactions on Visualization and Computer Graphics 23*, 6 (2017), 1576–1587. 1

[GH10] GOLDBERG J. H., HELFMAN J. I.: Comparing Information Graphics: A Critical Look at Eye Tracking. In *Proceedings of the 3rd BELIV'10 Workshop: BEyond Time and Errors: Novel evaLuation Methods for Information Visualization* (2010), BELIV '10, ACM, pp. 71–78. 1

[GWBI15] GOFFIN P., WILLETT W., BEZERIANOS A., ISENBERG P.: Exploring the Effect of Word-Scale Visualizations on Reading Behavior. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (2015), CHI EA '15, ACM, pp. 1827–1832. 1

[GWFI14] GOFFIN P., WILLETT W., FEKETE J.-D., ISENBERG P.: Exploring the Placement and Design of Word-Scale Visualizations. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 2291–2300. 1

[HNA*11] HOLMQVIST K., NYSTRÖM M., ANDERSSON R., DEWHURST R., JARODZKA H., DE WEIJER J.: *Eye tracking: A Comprehensive Guide to Methods and Measures*, 1 ed. OUP Oxford, 2011. 1

[KFBW16] KURZHALS K., FISHER B., BURCH M., WEISKOPF D.: Eye Tracking Evaluation of Visual Analytics. *Information Visualization 15*, 4 (2016), 340–358. 1

[PBLL12] PALMA J. P., BROWN P. J., LEHMANN C. U., LONGHURST C. A.: Neonatal Informatics: Optimizing Clinical Data Entry and Display. *NeoReviews 13*, 2 (2012), 81–85. 1

[PT94] POWSNER S., TUFTE E.: Graphical Summary of Patient Status. *The Lancet 344*, 8919 (1994), 386–389. 1

[PT13] PARSONS L. M., TINKELMAN D.: Testing the Feasibility of Small Multiples of Sparklines to Display Semimonthly Income Statement Data. *International Journal of Accounting Information Systems 14*, 1 (2013), 58–76. 1

[Tuf06] TUFTE E.: *Beautiful Evidence*, 1 ed. Graphics Press, 2006. 1