Indicator distance and color effects in comprehension of multiple time series graph

Rosemary Seva, Katrina Chinjen, Nicole Estoista, Judy Ann Wu
Industrial and Systems Engineering Department
De La Salle University
Manila, Philippines
rosemary.seva@dlsu.edu.ph, katrina_chinjen@dlsu.edu.ph, nicole_estoista@dlsu.edu.ph,
judy_wu@dlsu.edu.ph

Abstract

Line graphs are prevalently used to show trend especially time series. However, there is limited study on the influence of visual properties in complex comprehension tasks such as determining differences in slope. In this study, the effect of color hue and horizontal distance between points were considered in a slope task. Monochrome palettes using blue, green and red were developed to illustrate 4-time series data. Saturation was fixed while lightness were varied. Horizontal distance between points were set to 1.67, 2.00 and 2.5 cm while maintaining a constant height of 6 cm. Red resulted in the worst performance in terms of response time (RT) and accuracy. As for distance, data suggest that lower values negatively influence response time and accuracy. The best RT was obtained for high setting and highest accuracy for medium setting. Colors of monochrome palettes used did not significantly influence RT and accuracy. The hypothesis that red can improve perception of slope was rejected. In this study, graph understanding was improved although statistically insignificant by blue. High and medium horizontal distance settings significantly affected performance in slope tasks invalidating the hypothesis posited in this study. Accuracy can best be improved by using the medium distance setting while RT is best at high distance setting.

Keywords
data visualization, time series chart, slope task, color, distance

1. Introduction

Companies gather data that are sometimes too much to handle. The challenge is to organize and design visualizations that provide quick and accurate understanding of complex data (Ryan, 2015). People who have advance degrees find it difficult to interpret graphs (Bowen and Roth, 2005). It is the job of the graph designer to determine visual properties that will make the visualization comprehensible and attractive. Readability which is defined as the ability to extract information that is encoded in graphs is influenced by the color scheme (Thudt, 2007). Reading a graph presents numerous challenges that include slope interpretation (Janvier, 1998), detection of patterns (Bowen and Roth, 2005), and confusion due to visual clutter (Li and Moachieh, 2014). Comprehension, on the other hand, does not only include reading but also interpretation of graphs (Friel, et al. 2001).

The simplicity of the line graph lends itself to prevalent use in illustrating trends in multiple contexts. Time series data is usually represented using line graphs to show how a variable's behavior over time. However, the increase in the number of time series creates visual clutter that affects the readability of data. Visualization of multiple TS is facilitated by the use of colors and other unique visual properties such as line characteristics to distinguish each series of points Javed et al. (2010). However, there is limited research on the manipulation of graph design features on the readability of line graphs especially as information and task complexity increases. Cleveland and McGill (1984) noted the limitation of the human perceptual system in comparison of values on curved slopes which is part of graph interpretation (Cleveland and McGill, 1984).

Two visual properties of graphs: color hue and space were tackled in this study because they are relevant in improving the comprehension of multiple TS graphs (Javed et al., 2010). Seckler (2015) noted the dearth of literature on the effect of saturation on data visualizations independent from hue and brightness even in online environments.
The effect of horizontal distance between data points on graph comprehension is not well studied in literature. For line and bar graphs containing three indicators, response times were found to be sensitive to varying horizontal distances between data points (Gillan and Neary, 1992). The authors of this paper are not aware of other studies that considered horizontal distance between data points in line graphs especially multiple time series containing more than three data points. The complexity of this type of graph lends itself to misinterpretation and confusion because of information density. The design of multiple time series graphs is challenging because there are many data points oftentimes exceeding five to indicate trend. The design should facilitate accurate and quick interpretation especially in tasks that intermediate levels of understanding outlined by Glazer (2011). In this study, only the slope task was considered which requires identification of the time series with the highest increase throughout the entire time period (Beattie and Jones, 2002). This paper continues the work of Chinjen et al. (2018) on graph comprehension but in this paper only the slope task was considered in the analysis since it is the most complex task in interpreting graphs.

The study posits the following hypotheses:

H1. Use of red monochrome palette improves slope comparison of multiple TS graph

Comparing line graph slopes requires more information processing and time to view the graph. In order to identify the correct answer, the observer needs to focus on the detailed characteristics of the line. It is believed that for tasks that require close attention, the red hue will facilitate quick and accurate interpretation because the central part of the eye consists mostly of sensors that are sensitive to the red hue (Fortner and Meyer, 1997).

H2. As distance between indicators decreases, comprehension of 4-time series graphs improves

Gestalt principles suggest that distance (or proximity) provides visual aid during the extraction of potential trends in data (Shah and Freedman, 2011). However, this was only tested in the context of bar graphs. In addition, Heer et al. (2009) suggested that a large horizontal distance between points may negatively affect task performance. This study would therefore investigate whether decreasing the distance property may improve the user performance in readability of multiple time series line charts. Hence, being able to aid in extraction of trends would denote that completion time and accuracy relating to such tasks may be improved with decrease in distance.

2. Method

2.1 Participants

There are 88 university students that participated in the experiment which were qualified following a visual acuity and color vision test using the Snellen test and Ishihara color test, respectively. Only subjects with normal color vision were allowed to participate due to the influence of color on graph comprehension.

2.2 Apparatus

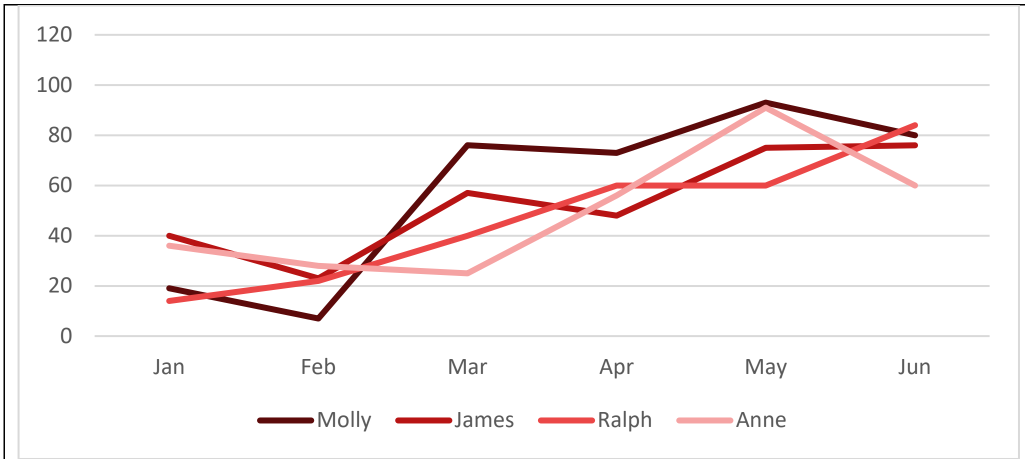
Variation of visual properties (color and distance), generation of the visualization, and the multiple-choice questionnaire were written in HTML. Experiments were conducted on a desktop computer monitor of 21.5” LED at 1920 X 1080 resolution, a standard mouse, and keyboard. Furthermore, each monitor was adjusted to the same level of contrasts and brightness.

2.3 Experimental conditions

External factors were controlled to avoid interference with the generated results. Experiments were conducted in a computer laboratory to ensure consistent ambient lighting as well as the condition and display size of the computer. Noise levels in the room during experimentation was kept at the minimum to avert any potential distractions. Furthermore, participants were allowed to adjust their seating and distance from the apparatus and make themselves comfortable.

2.4 Tasks

Participants of this study performed a slope task. An example of a slope task using the red monochrome palette is shown in Figure 1. During the experiment, participants were instructed not to compute for the slope but use their visual perception to judge the slope.



Which employee had the highest overall increase in sales?
☐ Molly
☐ James
☐ Ralph
☐ Anne

Figure 1. Sample Slope Task using Red Monochrome Palette

2.5 Variables

Independent variables

The variables manipulated in the experiments consist of the following:

1. Hue - different segments of the color wheel were considered to arrive at the base hues (red, green, and blue) which were developed into monochromatic palettes. The HSL (Hue-Saturation-Lightness) model was used because it is practical to use as it is already incorporated in most computer programming languages.

2. Horizontal Distance - refers to the space between two data points in the line graph. This space was altered to test the implication of a graph's size proportions to perception. A total of three conditions were tested for the distance experiment: low, medium, and high distance between data points. The low condition had a height to width ratio of 40%; the middle condition, 50% height to width ratio; and the high condition 60% height to width ratio. These settings were meant to represent long, average, and short graphs respectively. Those ratios were derived from the default ratios of the visualizations automatically generated by charting tools (e.g. MS Excel, ChartGo, etc.)

Dependent variables

Two performance metrics of graph comprehension were considered in the experiment as discussed below:

1. response time - generally defined as the time to respond to a stimulus recorded in time units (seconds). The interface developed recorded the response time. The recording of response time will begin once the data visualization along with the respective question is presented to the participant and then it will end once they have selected their answer.

2. accuracy - correctness of the participants' answers expressed as a percentage of total questions

2.6 Procedure

Development of monochrome palette

Red, green, and blue (RGB) hues were used in the experiment because they are approximately equally spaced in the color wheel and human beings possess three kinds of cones that are sensitive to short (445 nm), medium (535 nm), and long wavelengths (575 nm) of visible lights (Fortner and Meyer, 1997; Goldstein, 2007). The lightness and saturation of the selected base hues were then varied to maximize visual distinction (e.g. two colors of the same hues of green would be visually distinct due to the differences in their lightness or saturation).

The method proposed by Huang (2012) was used. A monochrome palette consisting of four variations of a single hue was generated as can be seen in Figure 2. First, the levels of saturation and lightness for the monochrome palette were derived by dividing the saturation and lightness scales into four approximately equal spaced levels: 20%, 40%, 60%, and 80%. Two sets of monochrome palettes were generated. For each set, either lightness or saturation was kept constant. Thus, there were 16 palettes developed for each base hue.

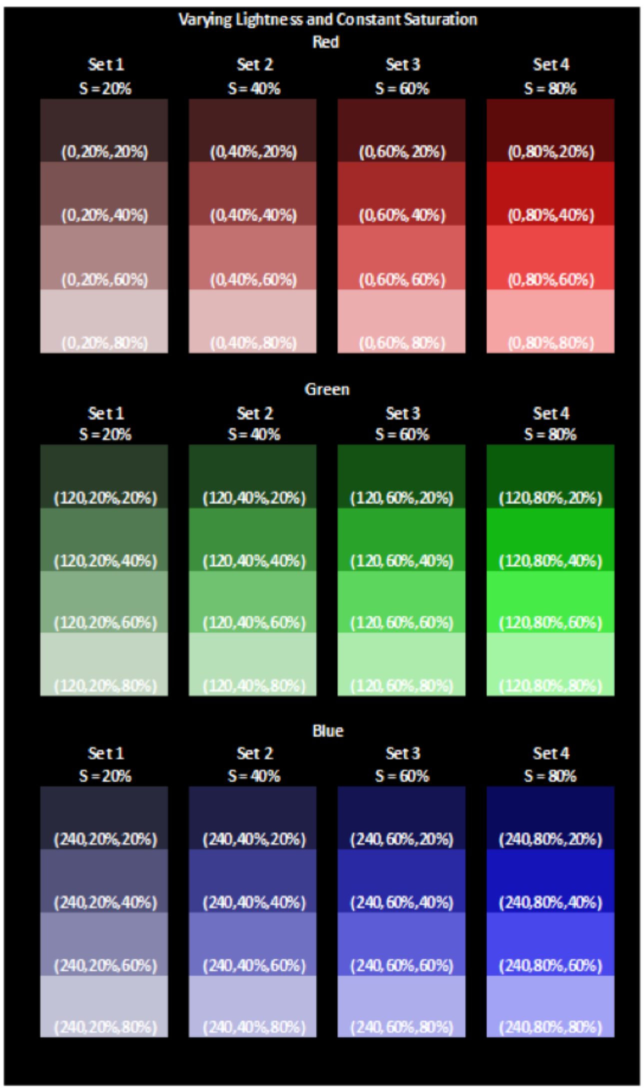


Figure 2. Monochrome Palettes for each Base Hue (RGB)

There are many versions of monochrome palettes generated. Contrast ratio was used to choose the palette to be used for the experiment since the color variation in a certain palette should be perceptible. The color set with the lowest standard deviation of contrast ratios were chosen for each hue. The mean contrast ratio for blue, green and red palette used were computed to be 3.56, 3.26, and 3.32, respectively. Color set 4 was used in the experiment.

Selection of distance settings

There are three conditions chosen for the distance experiment: low, medium, and high distance between data points. The height to width ratio of the low, medium and high settings were 40%, 50%, and 60%, respectively. These ratios were taken from default visualization ratios from known charting tools (e.g. MS Excel, ChartGo, etc.). Considering these ratios, the chart widths were varied while maintaining the same chart height. A chart height of 6 cm was chosen because this chart height can follow the specified height to width ratios by only adjusting the chart width. The specific settings are shown in Table 1. The color scheme of Excel was used to generate the graphs.

Table 1. Distance Settings

Distance between Data Points	**Total Chart Width**	**Height to Width Ratio**		
		40%	**50%**	**60%**
1.67 cm	10 cm	4.0 cm	5.0 cm	6.0 cm
2.00 cm	12 cm	4.8 cm	6.0 cm	7.2 cm
2.50 cm	15 cm	6.0 cm	7.5 cm	9.0 cm

2.7 Experimental Design

A within-subjects factorial design was used exposing each participant to 6 different conditions summarized below. Only one task was accomplished which is the slope task.

- Monochrome Palette (P): Blue, Green and Red
- Distance (D): 10 cm (low), 12 cm (medium), 15 cm (high)

2.8 Addressing Bias

Several biases can potentially affect the results of the experiment such as knowledge and ordering bias (Borgo et al., 2012), thus, a synthetic dataset was used to prevent knowledge bias. Ordering bias was minimized by randomizing the order of answer options and question presentation.

2.9 Experimental Procedure

The experiment was performed inside a room with a desktop computer. Participants were asked to sign consent forms and take the Ishihara Color Test and the Snellen Test. Participants with visual acuity problems were asked to wear their spectacles prior to proceeding for the actual experiment. Each participant was also advised to answer as quickly and accurately as possible.

Prior to the actual experiment, a tutorial was presented to instruct them how to answer using the program. A total of 6 questions were answered by each participant. Upon experiment completion, the participants underwent debriefing. After the experiment, participants were asked to answer a set of post-experiment questions. Each participant was compensated with a USD1 food voucher.

3. Results

Response time (RT) data violated the normality assumption based on its standardized residuals so non-parametric statistics was used in the analysis. Summary statistics obtained from the experiment are shown in Table 2. The blue monochrome palette had the best performance in terms of RT and comparable performance with green in terms of accuracy. Red resulted in the worst performance in terms of RT and accuracy. As for distance, data suggest that lower values negatively influence response time and accuracy. The best RT was obtained for high setting and highest accuracy for medium setting.

Table 2. Response Time Descriptive Statistics

		MEAN	SD	ACC
HUE	BLUE	12.81	5.30	63%
	GREEN	14.36	6.76	64%
	RED	14.44	7.34	57%
DIST	LOW	13.36	6.75	58%
	MED	11.94	5.54	81%
	HI	10.65	4.62	75%

Reaction time does not follow a normal distribution and is heavily skewed. As such, typical parametric tests such as Analysis of Variance (ANOVA) cannot be used to test the difference of mean RT for each type of palette. Kruskal-Wallis Test is the non-parametric counterpart of ANOVA and was conducted to determine if there are differences in RT according the hue of monochrome palette and distance. No significant differences (H=1.86, 2 d.f., p=0.395) were found in RTs for all hues but a statistically significant difference was found for distance (H=5.92, 2 d.f., p=0.05s).

Accuracy was analyzed using test of proportions. The hypothesis that the proportion of correct answers for different palettes and distances are the same. The results of the test of proportion are summarized in Table 3. Similar to RT, there were no significant differences in accuracy for hue unlike for distance. There was a very significant difference in accuracy between low and medium followed by low to high horizontal graph distance.

Table 3. Test of Proportion for Accuracy

HUE	z-value	p
Blue vs Green	-0.156	0.873
Blue vs Red	0.924	0.358
Green vs Red	0.768	0.441
DISTANCE		
Low vs Med	-3.269	0.001
Med vs High	0.908	0.363
Low vs High	-2.395	0.016

4. Discussion

This study considered a very specific task that is one of the most difficult in graph comprehension. It required a participant to visually compare different slopes from the multiple TS graph which results in visual confusion. McDermott et al. (1987) found that students are confused about differences in graph slope and height as well as interpreting changes. Even if students are good mathematically they cannot translate said skills in the process of interpreting graph situated in a particular context (Planinic et al., 2012). In this study, it was investigated if the use of monochrome palette can influence accuracy and RT which are crucial measures of comprehension. It is desired to know if the difficulty of the task can be mitigated by manipulating the visual properties of graphs.

Although the differences in RT and accuracy between the hues were not significant as can be seen in Tables 2 and 3 the blue monochrome palette performed slightly better than green and red as it has the smallest mean RT in Table 2. This means that participants are able to see the difference in slope when blue palette was used. It was expected that the abundance of red sensors in the eyes and the required focus inherent in the task will work to the advantage of this particular hue but results showed that the reverse happened. The relative responsiveness of the eyes to green and yellow green hue also did not contribute to quicker decision making or accuracy of answers (Fortner and Meyer, 1997). One possible explanation is the low contrast ratio of the red palette compared to blue. If the contrast ratio is low, it is more difficult to distinguish one line from another. It was clear that no specific hue has clear advantage over the another based on the insignificant result of the test on . Color was found to aid decision making in comprehending graphical reports especially if there is a time constraint (Benbasat and Dexter, 1986). However, the study of Benbasat and Dexter (1986) used only green without any variation in saturation and lightness. Thus, it is also possible that the use of different levels of saturation in constructing the monochrome palette could have caused confusion.

Contrary to color, horizontal distance between data points significantly influenced RT and accuracy. The hypothesis about distance posited in this paper was invalidated since high and medium distance settings resulted in best RT and accuracy, respectively. Such a result can be brought about by the nature of the slope task which can be difficult to perceive if the points are close to each other. A line graph that is too tall or narrow a disproportionate rate of change can be construed especially since it involves the angles and slopes of the lines. In contrast, if proportions are too short and wide, the rate of change will look small (Vlamis and Vlamis, 2015).

5. Conclusion

A specific graph comprehension task was considered in this study where hue and horizontal distance were considered as factors. Colors of monochrome palettes used did not significantly influence RT and accuracy. The hypothesis that red can improve perception of slope was rejected. In this study, graph understanding was improved although statistically insignificant by blue.

High and medium horizontal distance settings significantly affected performance in slope tasks invalidating the hypothesis posited in this study. Accuracy can best be improved by using the medium distance setting while RT is best at high distance setting.

The results of this study can be used to improve sales report presentation where audience are interested to know about trends indicated by slopes. Sales performance of different products over time are sometimes presented for comparison and eventual decision making. In presenting this kind of information, it is recommended that the distance between data points should be 2-2.5 cm with a height-to-width ratio of 50%.

References

Beattie, V., and Jones, M. J., The Impact of Graph Slope on Rate of Change Judgments in Corporate Reports, *Abacus,* vol. 38, no. 2, pp. 177-199, 2002.

Benbasat, I., and Dexter, A. S., An Investigation of the Effectiveness of Color and Graphical Information Presentation under Varying Time Constraints. *MIS Quarterly,* vol. 10, no. 1, pp. 59-83, 1986.

Borgo, R., Abdul-Rahman, A., Mohamed, F., Grant, P. W., Reppa, I., Floridi, L., and Chen, M., An Empirical Study on Using Visual Embellishments in Visualization. *IEEE Trans Vis Comput Graph,* vol. 18, no. 12, pp. 2759-2768, 2012.

Bowen, G. M., and Roth, W.-M., Data and graph interpretation practices among preservice science teachers. *Journal of Research in Science Teaching,* vol. 42, no. 10, 1063-1088, 2005.

Chinjen, K., Estoista, N., Seva, R., and Wu, J. A. (2018). *The effect of color and space between graph elements on comprehension*, Proceedings of the 5th Southeast Asian Ergonomics Conference, Bangkok, Thailand, December 12-14, 2018.

Cleveland, W. S., and McGill, R., Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, vol. 79, no. 387, pp. 531-554, 1984.

Fortner, B., and Meyer, T. E., *Number by colors a guide to using color to understand technical data*, Springer-Verlag, New York, 1997.

Friel, S.N., Curcio, F. R., and Bright, G.W., Making sense of graphs: critical factors influencing comprehension and instructional implications, *Journal for Research in Mathematics Education*, vol. 32, no. 2, pp. 124-158, 2001.

Glazer, N., Challenges with graph interpretation: a review of the literature, *Studies in Science Education*, vol. 47, no. 2, pp. 183-210, 2011.

Goldstein, E. B., *Sensation and Perception*, 7th edition, Wadsworth, Belmont, CA, 2007.
Huang, S. M.. The rating consistency of aesthetic preferences for icon-background color combinations. *Applied Ergonomics,* vol. 43, no. 1, pp. 141-150, 2012.
Javed, W., McDonnel, B., and Elmqvist, N., Graphical Perception of Multiple Time Series, *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 6, pp. 927-934, 2010.
McDermott, L. C., Rosenquist, M. L., & van Zee, E. H. Student difficulties in connecting graphs and physics: Examples from kinematics, *American Journal of Physics,* vol. 55, no. 6, pp. 503-513, 1987.
Planinic, M., Milin Sipus, Z., Katic, H., Susac, A., and Ivanjek, L., Comparing student understanding of line graph slope in physics and mathematics. *International Journal of Science and Mathematics Education,* vol. 10, no. 6, 2012.
Vlamis, D., and Vlamis, T., *Data Visualization for Oracle Business Intelligence 11g*, McGraw-Hill Education Group, USA, 2015.

Biographies

Rosemary Seva is a Professor of Industrial Engineering and former Dean at the Gokongwei College of Engineering - De La Salle University, Philippines. She is the current President of the Human Factors and Ergonomics Society of the Philippines (HFESP) and the past President of the Southeast Asian Network of Ergonomics Societies (SEANES). She is one of the Top 150 scholars and scientists in the Philippines from Webometrics. She obtained her doctorate degree at the Nanyang Technological University (Singapore) where she wrote her dissertation on affective design of mobile phones. She has a master's degree in Ergonomics from the University of New South Wales (Australia) and another master's degree in Industrial Engineering from De La Salle University. She has written a number of papers that dealt on usability, product design, cognitive task analysis, and physical ergonomics.

Katrina Chinjen is a Production Supervisor at LF Logistics where she has participated in projects involving different manufacturing clients. She earned a B.S. in Industrial Engineering at De La Salle University - Manila, Philippines where she and her team graduated with the Most Outstanding Thesis Award of their course. Her field interests include productivity, digitalization, and cost efficiency.

Nicole Anne Therese Estoista is a software engineer. She earned a BS in Industrial Management Engineering with a Minor in Information Technology and graduated with the Most Outstanding Thesis Award from De La Salle University.

Judy Ann Wu is a country planning and material controller in a telecommunications company. She earned her B.Sc in Industrial Engineering at De La Salle University and graduated with Most Outstanding Thesis Award together with her team. Her field interests include applications of machine learning and data science in supply chain optimization.