Tranquille, Dominique, Stapleton, Gem, Burton, Jim and Rodgers, Peter (2016) *Evaluating the Effects of Colour in LineSets.* In: Lecture Notes in Computer Science. 9th International Conference on Diagrammatic Representation and Reasoning (Diagrams 2016). 9781. pp. 283-285. Springer-Verlag Berlin ISBN 978-3-319-42332-6. E-ISBN 978-3-319-42333-3.

Evaluating the Effects of Colour in LineSets

Dominique Tranquille[1], Gem Stapleton[1], Jim Burton[1], and Peter Rodgers[2]

[1]Visual Modelling Group, University of Brighton, UK
d.tranquille1@uni.brighton.ac.uk
{g.e.stapleton,j.burton}@brighton.ac.uk
[2]University of Kent, UK
p.j.rodgers@kent.ac.uk

1 Introduction

Data items often lie in overlapping sets and a number of set visualization techniques have been developed in recent years [1, 4]. An example is in Figure 1, which visualizes sets by overlaying lines on an existing visualization of data items. As LineSets are an overlay technique, they can be applied to many different data sets. Here, we apply them to network diagrams. LineSets are composed of lines that are overlaid on nodes or a node-link network; examples are given in Figures 1 and 2. The set-lines are labelled to indicate the represented set. Nodes represent individual entities (data items). A node represents a member of a set if the set-line passes through it. Nodes that are not passed though by set-lines are not members of any of the sets. Two nodes represent related data items if an edge passes between them. Alper et al.'s initial paper on LineSets focussed on exploring the potential of their new technique [1]. Their research established that Linesets should be generated with paths that are as linear as possible as well as being smooth. However, there are a number of other graphical choices to be made when drawing LineSets; one of these choices is colour. This paper identifies how colour (hue, value, or monochrome) should be applied to LineSets drawn on networks. The study materials and collected data is available at `http://www.cem.brighton.ac.uk/research/VMG/linesets-study-2015`.

2 Experiment Design

The aim was to establish how different colour treatments applied to LineSets affect task performance. The three colour treatments that we used were unique hues, varying values and monochrome. For our unique hues colour treatment, we selected colours that were uniformly distinct from each other. The colour palette we created for the varying values treatment consisted of a single colour hue that was broken down into stepped levels of lightness (Figure 2). The colours were generated from Color Brewer [3]. Black was used for all items in the monochrome treatment.

We used a between-group design with repeated measures and three groups. We recorded two dependent variables: the time taken to answer the question and

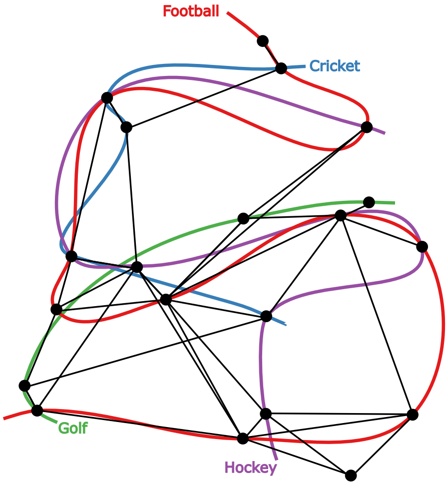

Fig. 1. Unique Hues.

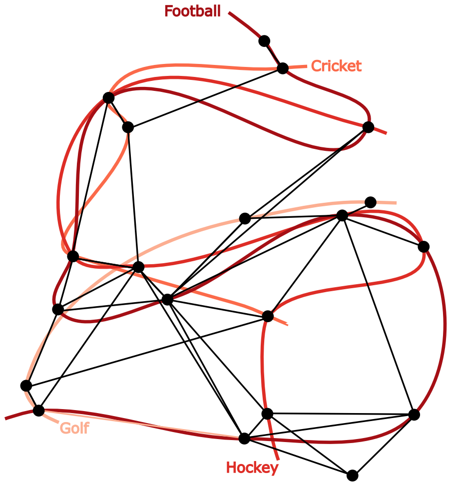

Fig. 2. Varying Values.

whether the answer was correct. If colour treatment impacts on task performance then we expect a significant difference to exist in either accuracy or time. All of the tasks performed by participants in our study fit into Simonetto et al.'s group-level graph visualization taxonomy [5]. This taxonomy identifies four question types suited to LineSets drawn on networks: group only (about sets), group-node (about sets and data items), group-link (about sets and network connections) and group-network (about sets, data items and network connections). Covering all four task types allowed us to collect meaningful data about general task performance.

3 Statistical Analysis

We used the data collected from 60 participants together with the six pilot participants, as no changes had been made to the experiment. This gave us a total of 66 participants (55 M, 11 F, age: average 24, range 18 to 46). Each participant answered 24 questions giving us a total of 1584 observations. There were a total of 31 unanswered 'timeouts', where participants had not provided an answer within the two minutes allowed (monochrome 19, hues 6 and values 6). These were removed from our data prior to analysis, as no answer to the question was provided.

Of the 1553 questions for which answers were provided within the 2 minutes allowed, there were 556 errors (35.8%). Monochrome had the highest inaccuracy rate with 258 errors (50.7%). Varying hues and values had inaccuracy rates of 138 (26.4%) and 160 (30.6%) errors respectively. We conducted a chi-square test, giving a p-value of 0.000. It was established that monochrome yielded significant more errors than both varying hues and varying values. However, varying hues and varying values were not significantly different.

With regards to time taken, the overall mean completion time, for the 997 correct answers, was 38.01 seconds and the standard deviation was 22.14. Varying hues achieved the lowest mean completion time and standard deviation of 30.92 (18.07). Varying values and monochrome achieved mean completion times and standard deviations of 37.36 (20.91) and 49.81 (24.57) respectively. We performed an ANOVA test in order to determine the overall affect of colour treatment on time performance. However, the data were not normally distributed so we applied a log transformation which yielded a skewness of -0.01 and, thus, rendered our data suitable for analysis. The results revealed significant differences between the colour treatments, with a p-value of 0.000. A Tukey Test revealed that varying hues allowed participants to perform significantly faster than varying values. Both these treatments were significantly faster than monochrome.

4 Conclusion and Future Work

This paper set out to address the question *how does the use of colour affect the comprehension of LineSets?* We selected three colour treatments based on perceptual theories on the use of colour: monochrome, varying values and varying hues. Taking into account both our error analysis and time analysis, LineSets should be treated with varying hues.

Whilst we have focused on the use of colour in LineSets, there are other graphical properties that need to be understood. Properties such as size and shape, along with colour, can have a profound impact on perception [2]. In the context of LineSets, size corresponds to graphical properties such as set-line thickness and node size. Regarding shape, the set-lines can take many routes when being overlaid on nodes, and be either smooth or composed of a series of straight line segments. Evaluating these, and other, graphical properties will lead to a more comprehensive understanding of how to draw LineSets to aid task performance.

References

1. Alper, B., Riche, N.H., Ramos, G., Czerwinski, M.: Design study of LineSets, a novel set visualization technique. IEEE Transactions on Visualization and Computer Graphics 17(12), 2259–2267 (2011)
2. Bertin, J.: Semiology of Graphics. University of Wisconsin Press (1983)
3. Brewer, C., Harrower, M., B., D.H., Sheesley, Woodruff, A.: Color brewer (2013), `http://colorbrewer2.org/`, accessed 14 August 2015
4. Meulemans, W., Riche, N.H., Speckmann, B., Alper, B., Dwyer, T.: Kelpfusion: A hybrid set visualization technique. IEEE Transactions On Visualization And Computer Graphics 19(11), 1846–1859 (2013)
5. Saket, B., Simonetto, P., Kobourov, S.G.: Group-level graph visualization taxonomy. In: EuroVis - Short Papers (2014)