EuroVis 2022 Short Paper

ISBN 978-3-03868-184-7

magentahttps://doi.org/10.2312/evs.20221089

The Effect of Graph Layout on the Perception of Graph Density: An Empirical Study

Elektra Kypridemou *[1,2], Michele Zito[1], and Marco Bertamini[2]

[1] *Department of Computer Science, University of Liverpool, UK*
[2] *Visual Perception Lab, Department of Psychological Sciences, University of Liverpool, UK*

Abstract

The visual representation of a graph is crucial in understanding and analyzing its properties. In this empirical study, we examine the effect of different drawing layouts on our perception of graph density. We treat density as an absolute property of the graph and use a Yes-No design, where participants have to decide whether a graph has a given density or not. We compare a simple grid layout with well-known planar and spring layouts. We also introduce an alternative 'improved' grid layout, which reduces the number of crossings while keeping most of the simplicity of the original grid layout. Results show that our 'improved' version of the grid layout facilitated performance on the task, compared to the original one. Moreover, participants were biased into judging graphs as denser when drawn with the original grid layout, while tended to perceive graphs as less dense when drawn with the planar and grid layouts. In contrast to previous studies on graph density perception, this is the first indication that the chosen layout can influence our perception of the graph's density.

1 Introduction

The way humans perceive the density of elements within a scene has been long studied within the literature [10, 11]. Previous studies examined the relationship between numerosity and density discrimination, as well as the underlying mechanisms of the two processes [6, 1, 2, 25]. In this study, we focus our attention on the perceived density of graphs, when they are depicted as node-link diagrams.

The way visually represent data can strongly influence how the viewers perceive the data [26, 27, 22, 20]. The purpose of this study is to examine how the way we choose to depict a graph on the plane (layout) affects our perception of the graph's density, as a graph property. In other words, we study how the visual

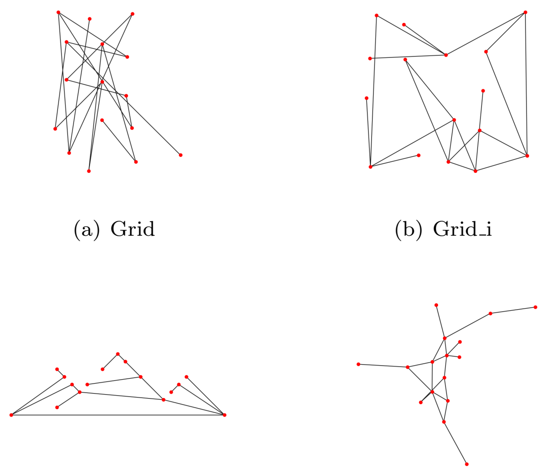
*Corresponding author: Elektra Kypridemou, e.kypridemou@liverpool.ac.uk

density of the layout affects our judgements about the subjective density of a graph. Graph density is a metric that indicates how 'full' a graph is by describing the relationship between the number of edges and the maximum number of edges, as a function of the number of vertices of a graph. Although a graph's density affects its drawing, it's not identical to the texture density of its drawing.

Previous empirical studies towards this direction, have examined the effect of graph layouts on the perception of a number of graph properties, including density. In a first experiment to model human ability to discriminate graph properties, Soni et al. [23] used the Just Noticeable Difference (JND) methodology to examine the effect of three different layouts (Force Directed, Circular, MDS) on the perceived density of graphs of 100 nodes each. Results found no significant difference on the perception of density between the three layouts. In a later study, Kypridemou et al. [19], examined the effect of four different layouts (Circular, Grid, Planar, and Spring) on the perception of density and other properties on smaller graphs of 16 nodes. Using the two-alternative forced choice (2AFC) methodology, drawings of the same layout were displayed side by side and participants had to decide which of the two graphs had the respective property. For the density task, stimuli were planar connected graphs of 16 nodes, with either 16 or 18 edges each. Replicating Soni et al.'s results on smaller graphs, none of the four layouts was found to facilitate performance more than others.

In this study, we treat graph density as an absolute property of the graph (i.e., a graph having 20 edges), rather than a relative judgment based on the density of another graph (i.e., a graph being denser than another graph). This way, we extend the previous work of Kypridemou et al. [19], by transforming the task from a comparison to the detection of a property. Hence, instead of presenting two stimuli at a time (target and non-target), we choose to use a Yes-No answer methodology and present a single stimulus per trial. The Yes-No design also allows control over the amount of time that participants spend on each stimulus, which would not be possible with the 2AFC method. Moreover, the new design makes possible further analyses on the effect of particular graph aesthetics (i.e., features of the drawings) on the perceived density. This allows to draw conclusions about which features of the stimuli affect the perceived density. The 2AFC design provides no insights about which of the two stimuli (left or right) primarily affected the human decision in each trial, and therefore such analyses would not be possible. Finally, in Kypridemou et al. [19], in both the Circular and the Grid layouts, nodes were arbitrarily placed on pre-defined positions. To avoid having two similar baseline layouts, we replace the Circular layout by an improved version of the Grid one.

The aim of this experiment is to examine the effect of different graph layouts on the perception of graph density, using a mixture of qualitative and quantitative methods. In the following sections we present our methodology and procedure, we describe the stimuli, and we discuss the results of the experiment.



(a) Grid (b) Grid_i

(c) Planar (d) Spring

Figure 1: Exemplar target (20 edges) and non-target (16 edges) stimuli of the experiment, for different layouts.

2 Method

To define the presentation time per trial and the difference between target and non-target stimuli, we ran a pilot study on three participants. The goal was to make the task challenging enough, but also avoid close to random performance. Because we presented a single stimulus at a time instead of two, we considered decreasing the presentation time compared to the previous study. Moreover, because of the Yes-No design, we also considered increasing the difference in density between target and non-target graphs. Based on the pilot study's results, we decided to decrease the presentation time from 3000 to 2000 msec, and use target and non-target graphs with a difference of four edges (16 versus 20 edges) instead of two. Moreover, to explore differences in layouts beyond the % correct answers, in this study we use signal detection theory [24] to analyse the d prime (d') and bias (c) dependent variables. To get adequate data for this analysis, we increase the number of trials to 200 per participant.

Based on the above decisions about the experiment's parameters, we concluded the experimental design. The participants viewed drawings of graphs of 16 nodes, with either high (20 edges) or lower (16 edges) density, and had to decide whether the graph had 20 edges (target graph). The graphs were drawn with each of the four layouts of Figure 1, and were presented one in each trial. As shown in Figure 2, each trial started with a fixation cross displayed at the centre of the screen for 1500 msec, after which one of the stimuli appeared for 2000 msec. Afterwards, a prompt message remained on the screen until the participants provided their response using the keyboard. The experiment consisted of 200 trials, which appeared on a fully random order.

2.1 Participants and Procedure

Sixteen participants (7 male, 9 female) aged between 18 and 41 ($M = 24.44, SD = 8.59$) participated in the study. All of them reported normal or corrected to normal vision and had no previous experience in graphs. This experiment passed the local ethics committee approval and took place at the Visual Perception Lab of the Psychology Department of the University of Liverpool.
The experimenter explained to the participants that they are going to see graphs of 16 nodes, with either 16 edges or 20 edges, and they will have to decide whether the graph has 20 edges. Then, the participants had the chance to view eight exemplar stimuli (4 from each layout: 2 target, 2 non-target), guess the correct answer, get feedback, and discuss their answers with the experimenter. After the instructions, participants reported their confidence level on understanding the task on a 4-level Likert scale. Then, they moved into a quiet darkened room and used a chin rest of adjustable height, to control the distance between their head and the monitor. Participants performed a training phase of 16 trials, during which they received auditory feedback for each trial, and then they did the actual experiment, which consisted of 200 trials without feedback. Every 50-trial sessions, a message appeared on the screen informing the participants about their progress so far. After the experiment, they reported in a written format the strategies they used while performing the task. This post-experiment question helped us gain some insights about the internal processes and mechanisms, and identify specific features of the stimuli that might have affected their performance. The average time of the experiment per participant, excluding the instructions and training phase, was 9: 30 minutes ($M = 9.5, SD = 1.13$ minutes). In the following section, we describe how we generated the graphs, applied the different graph drawing algorithms, and produced the final images of the stimuli. When necessary, we provide the definitions of graph-theoretic terms on footnotes.

2.2 Stimuli

All graphs, drawings, and images were generated using Python 3.8 and NetworkX 2.5 [15]. The experiment was coded using the PsychoPy 2021.2 library [15] and Python 3.6. All stimuli were drawings of simple [1] graphs of 16 nodes each, drawn with one of the four layouts shown in Figure 1. Target stimuli had 20 edges, while non-target had 16 edges. To limit any effects of confounding variables we also ensured that both target and non-target stimuli were planar [2]

[1] Simple graphs are unweighted, undirected graphs containing no edge loops (i.e., edges that connect a node to itself) or parallel edges (i.e., edges incident to the same two nodes).

[2] Planar graphs are those that *can be* drawn on the plane in such a way that no edges cross each other. However, only the planar graph drawing algorithms guarantee that planar graphs

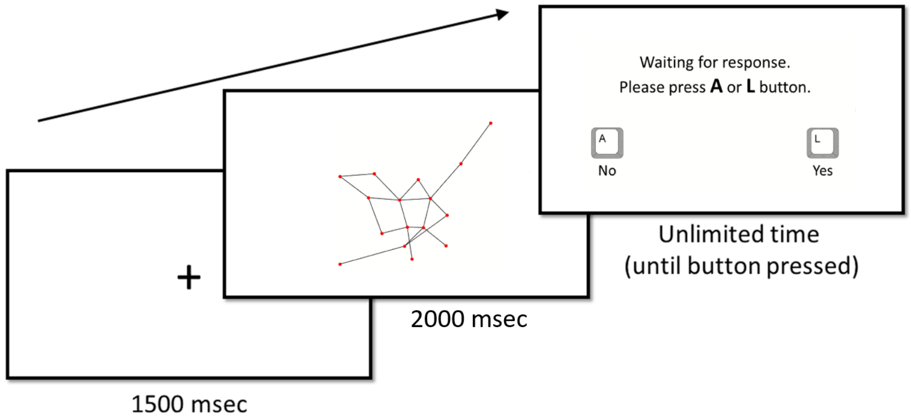
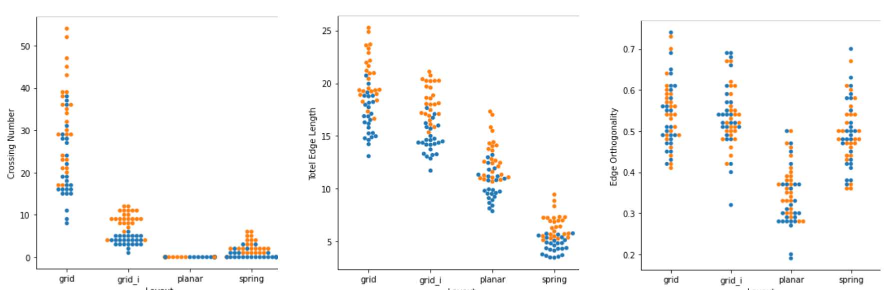


Figure 2: Event sequence of one trial. The participants had to respond to the question ‘Does the graph have 20 edges?’.



(a) Crossing Number (b) Total Edge Length (c) Edge Orthogonality

Figure 3: Aesthetics of the 200 stimuli of the experiment. Orange and blue dots represent graphs of 20 and 16 edges, respectively.

and connected [3].

We generated the non-target graphs using the $G_{n,m}$ random graph model by *Erdős-Rényi* [12] (with $n = 16$ nodes and $m = 16$ edges), discarding all resulting graphs that were not both connected and planar. To generate the target graphs, we added to each of the non-target graphs four more edges at random, using rejection sampling to avoid adding loops and parallel edges, while ensuring that the resulting graphs were also planar.

After generating the graphs (25 target, 25 non-target), we drew each one of them using the four layouts of Figure 1. For the *Grid* layout (Figure 1.a), each node was arbitrarily placed on a rectangular 5×5 grid with spacing d. The grid placement of nodes induces a high chance that edges overlap, which may introduce visual ambiguities [3, 14]. We resolved this by adjusting the grid positions in a disk of radius of $0.28 \cdot d$, centered on their original position of the grid points.

The *Improved Grid (Grid_i)* layout (Figure 1.b), was a variation of the previous algorithm, which was optimized in terms of some of the most common graph drawing aesthetics [21, 8, 9]. Starting from an initial placement of the Grid layout, the algorithm computes a total aesthetic metric $\mathcal{M}$ as a linear combination of the normalised planarity (np) [4], the average edge length (ael), the edge orthogonality (eo) [5], and the weighted average shortest path length ($aspl$) [6], using the edge lengths as weights. The output of the algorithm is the placement that optimizes $\mathcal{M}$ over 1000 iterations, while prioritizing the minimization of the number of crossings over the other aesthetics. Figure 3 shows the differences in three of the aesthetics before (Grid) and after (Grid_i) the optimization was applied. The Improved Grid layouts were significantly improved compared to the Grid layouts in all chosen aesthetics, except eo. The Grid_i layouts had 6.42 crossings on average, with $M(np) = 0.95$, $M(ael) = 0.92$, $M(eo) = 0.53$, $M(aspl) = 2.93$, while the Grid layouts of the same graphs had 26.8 crossings on average, with $M(np) = 0.78$, $M(ael) = 1.04$, $M(eo) = 0.54$, $M(aspl) = 3.27$.

Planar layouts (Figure 1.c) are drawings in which no edges cross each other ($np = 1$). Since we had generated our graphs to be planar, we could apply a planar layout to draw them in such a way. We used the Left-Right Planarity Test [7] and the Chrobak and Payne drawing algorithm [5] to get a straight-line planar drawing on a $(2n - 4) \times (n - 2)$ grid, where n is the number of nodes of the input graph. Since the Chrobak and Payne algorithm is using a preliminary

will be drawn with zero edge crossings.

[3] A graph is said to be connected if for any pair of nodes there is a path connecting them.

[4] The normalized planarity, or else 'normalized edge crossing metric', is the actual number of crossings in the drawing scaled against the maximum possible number of crossings. It varies between 0 and 1, with values closer to 1 indicating a good quality drawing. See [21] for definition.

[5] The edge orthogonality indicates the extend to which nodes and edges of the drawing follow the points and lines of an imaginary Cartesian grid. It varies between 0 and 1, with 1 indicating a good quality drawing [21].

[6] The average shortest path length is the average number of steps along the shortest paths for all possible pairs. For the weighted version, the shortest path is the one that minimizes the sum of the respective edge weights.

triangulation step, the resulting layout has a triangular-shaped outline that is internally triangulated.
Spring layouts (Figure 1.d) are generated by a type of Force-directed (FD) algorithms that model edges as mechanical springs and vertices as charged particles. The model applies forces to the nodes based on their relative positions in space, and simulates their motion until they reach an equilibrium that minimizes the total energy of the system. FD layouts are well known for producing aesthetically pleasing drawings that tend to exhibit synmmetries, uniform node distribution and few edge crossings [18]. We used the Fruchterman-Reingold algorithm [13], a traditional FD algorithm in which edges are simulating springs of fixed length. As a result, connected nodes are drawn near to each other, but not too close to each other, resulting to more uniform edge lengths [17, 4].
The resulting 200 drawings were depicted as node-link diagrams of red dots of fixed size (20 pixels) and black lines of fixed thickness (1.0 pixel). We also ensured that all drawings were enclosed on an imaginary fixed disk, centered on the canvas area, and were occupying a relatively equivalent visual area. All stimuli were saved as images of 180 dpi and 1152×864 pixels. They were presented on a 15.6" LCD monitor with a 60 Hz refresh rate, and a 1920×1080 pixels resolution, at a fixed viewing distance of 57 cm.

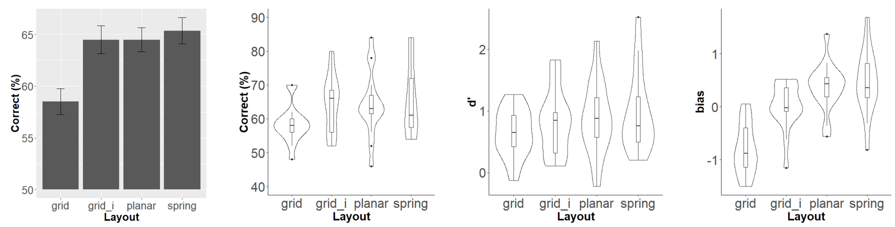
3 Results

3.1 Questionnaire Results

Prior to performing the experiment, the majority of participants (15 of 16) reported 'very high confidence' on their understanding of the task, with only one reporting having 'high confidence'. After the experiment, 12 out of 16 participants reported that they had used a specific strategy while performing the task. Four of the participants mentioned that the existence of many high-degree nodes indicated a denser graph. The number of closed cycles or triangles was another prominent feature that seemed to affect their judgements. Two of the participants reported using the number of cycles of the shape as an indication of its density, while one of them specifically reported that the existence of more than one cycles was a sign of target stimuli. The edge lengths also guided participants' judgements about density, with shorter lines indicating higher density. Other answers included the existence of crossing lines, and the complexity of the shapes, with simpler shapes indicating less dense graphs. One participant reported that when lines were closer to each other it felt like the graph was denser, but also observed that maybe that was a visual illusion.

3.2 Quantitative Results

To check whether overall performance was better than random answering, we performed an one-sample t-test with a 99% confidence interval. The mean % correct score was found to be significantly different than 50% for the whole



(a) % correct (b) % correct (c) sensitivity (d') (d) bias (c)

Figure 4: Bar plot of % correct and violin plots of all performance metrics (% correct, sensitivity, bias) per layout. Error bars refer to ±1 SEM.

task ($t(63) = 11.73, p < 0.001$). After applying a Bonferroni correction, to account for multiple hypothesis testing, we found that performance was also better than chance for each layout separately ($t(15) = 6.04, p < 0.001$ for Grid, $t(15) = 6.87, p < 0.001$ for Grid_i, $t(15) = 6.01, p < 0.001$ for Planar, and $t(15) = 5.79, p < 0.001$ for Spring).

We also calculated the hit (H) and false alarm (FA) scores, as well as the sensitivity (d') and bias (c), for each layout for each participant. To correct any H and FA scores of 1 or 0, we applied the *log-linear rule* [16]. The results of all performance scores per layout are shown in Figure 4. There was a significant main effect of layout on the % correct performance ($F(3, 45) = 6.33, p < 0.01$), with the Grid layout having significantly lower % correct performance than the rest, but still higher than chance. However, the d' measure of sensitivity was not significantly different for any of the layouts ($F(3, 45) = 2.19, p = 0.10$), indicating that the layout did not affect participants' sensitivity on detecting the target graphs. Finally, there was a significant main effect of the layout on the bias metric c ($F(3, 45) = 21.27, p < 0.001$). The bias of the Grid layout was clearly below zero ($M(c) = -0.81$), indicating the tendency of participants to reply 'yes' on graphs drawn with it. On the other hand, for the Planar and Spring layouts bias was found to be positive ($M(c) = 0.38$, $M(c) = 0.47$ respectively). The Improved Grid layout did not bias participants towards any direction ($M(c) = -0.016$).

4 Discussion

Unlike previous findings on graphs of similar and larger size [23, 19], we found a significant main effect of layout on the % correct performance of density judgements. We also found a significantly different behaviour between layouts when analysed the bias metric. When participants saw a Grid drawing, they had the tendency to perceive it as a high-density graph of 20 nodes, regardless of the subjective density of the graph. On the other hand, for the Planar and Spring layouts, they tended to perceive both target and non-target graphs as low-density graphs. This is particularly interesting, when we consider the aesthetics metrics of the different layouts. Grid drawings tend to have considerably more crossings and longer edges than the Planar and Spring layouts (see Figure 3). Hence, the quantitative results on bias are consistent with the strategies that participants reported in the online questionnaire. The higher number of crossings and the more 'complicated' drawings were some of the features that they mentioned as indicators of high-density graphs. The results indicate that specific aesthetics of the drawings, such as the crossings number and the total edge length, seem to be good predictors for the perceived density of a graph.

5 Conclusion

Previous studies in small [19] and larger [23] graphs, found no significant effect of the graph's layout on our perception of its density. In this study, we analyse the participants' sensitivity and bias in detecting a given density. In contrast to the two previous studies, this is the first indication that the graph layout might affect the perceived density of a graph. Results show that specific features of the layouts might bias our judgements about graph density. Drawings that seem to be more complex and unstructured (e.g., have more crossings) tend to bias us towards overestimating the graph's density, while more aesthetically pleasing and readable drawings (e.g., fewer crossings, more uniform edge lengths) tend to make the graphs look less dense.

This was a first indication that the layout, as well as the specific aesthetics of a drawing, might bias our perception of graph density. However, our investigation was limited to graphs of specific size and density. To fully explore the human perception of graph density, more extensive work needs to be done towards this direction. Future studies can explore a larger range of values regarding the size of the graphs, as well as the differences in number of edges between target and non-target graphs.

References

[1] Giovanni Anobile, Guido Marco Cicchini, and David C Burr. Separate mechanisms for perception of numerosity and density. *Psychological science*, 25(1):265–270, 2014.

[2] Giovanni Anobile, Marco Turi, Guido Marco Cicchini, and David C Burr. Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *Journal of Vision*, 15(5):4–4, 2015.

[3] Chris Bennett, Jody Ryall, Leo Spalteholz, and Amy Gooch. The aesthetics of graph visualization. *Computational aesthetics*, 2007:57–64, 2007.

[4] Franz J Brandenburg, Michael Himsolt, and Christoph Rohrer. An experimental comparison of force-directed and randomized graph drawing algorithms. In *International Symposium on Graph Drawing*, pages 76–87. Springer, 1995.

[5] Marek Chrobak and Thomas H Payne. A linear-time algorithm for drawing a planar graph on a grid. *Information Processing Letters*, 54(4):241–246, 1995.

[6] Steven C Dakin, Marc S Tibber, John A Greenwood, Michael J Morgan, et al. A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, 108(49):19552–19557, 2011.

[7] Hubert de Fraysseix and Pierre Rosenstiehl. A characterization of planar graphs by trémaux orders. *Combinatorica*, 5(2):127–135, 1985.

[8] Cody Dunne, Steven I Ross, Ben Shneiderman, and Mauro Martino. Readability metric feedback for aiding node-link visualization designers. *IBM Journal of Research and Development*, 59(2/3):14–1, 2015.

[9] Cody Dunne and Ben Shneiderman. Improving graph drawing readability by incorporating readability metrics: A software tool for network analysts. *University of Maryland, HCIL Tech Report HCIL-2009-13*, 2009.

[10] Frank H Durgin. Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human perception and performance*, 21(1):149, 1995.

[11] Frank H Durgin. Texture density adaptation and visual number revisited. *Current Biology*, 18(18):R855–R856, 2008.

[12] Paul Erdős and Alfréd Rényi. On random graphs i. *Publicationes Mathematicae*, 6:290–297, 1959.

[13] Thomas MJ Fruchterman and Edward M Reingold. Graph drawing by force-directed placement. *Software: Practice and experience*, 21(11):1129–1164, 1991.

[14] Helen Gibson, Joe Faith, and Paul Vickers. A survey of two-dimensional graph layout techniques for information visualisation. *Information visualization*, 12(3-4):324–357, 2013.

[15] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[16] Michael J Hautus. Corrections for extreme proportions and their biasing effects on estimated values ofd. *Behavior Research Methods, Instruments, & Computers*, 27(1):46–51, 1995.

[17] Jie Hua, Mao Lin Huang, and Guohua Wang. Graph layout performance comparisons of force-directed algorithms. *International Journal of Performability Engineering*, 2018.

[18] Stephen G Kobourov. Spring embedders and force directed graph drawing algorithms. *arXiv preprint arXiv:1201.3011*, 2012.

[19] Elektra Kypridemou, Michele Zito, and Marco Bertamini. The effect of graph layout on the perception of graph properties. *EuroVis (Short Papers)*, pages 1–5, 2020.

[20] Brian Ondov, Nicole Jardine, Niklas Elmqvist, and Steven Franconeri. Face to face: Evaluating visual comparison. *IEEE transactions on visualization and computer graphics*, 25(1):861–871, 2018.

[21] Helen C Purchase. Metrics for graph drawing aesthetics. *Journal of Visual Languages & Computing*, 13(5):501–516, 2002.

[22] Aidan Slingsby, Jason Dykes, and Jo Wood. Configuring hierarchical layouts to address research questions. *IEEE transactions on visualization and computer graphics*, 15(6):977–984, 2009.

[23] Utkarsh Soni, Yafeng Lu, Brett Hansen, Helen C Purchase, Stephen Kobourov, and Ross Maciejewski. The perception of graph properties in graph layouts. In *Computer Graphics Forum*, volume 37, pages 169–181. Wiley Online Library, 2018.

[24] Harold Stanislaw and Natasha Todorov. Calculation of signal detection theory measures. *Behavior research methods, instruments, & computers*, 31(1):137–149, 1999.

[25] Alberto Testolin, Serena Dolfi, Mathijs Rochus, and Marco Zorzi. Visual sense of number vs. sense of magnitude in humans and machines. *Scientific reports*, 10(1):1–13, 2020.

[26] Barbara Tversky. Visualizing thought. In *Handbook of human centric visualization*, pages 3–40. Springer, 2014.

[27] Cindy Xiong, Vidya Setlur, Benjamin Bach, Eunyee Koh, Kylie Lin, and Steven Franconeri. Visual arrangements of bar charts influence comparisons in viewer takeaways. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):955–965, 2021.