

The Effect of Visual Mark Arrangement on Value Estimation Accuracy in Waffle Charts

Dheerissara Lertpiriyasap*
Chulalongkorn University

Puripant Ruchikachorn†
Chulalongkorn University

ABSTRACT

This study investigates whether visual mark arrangement, color ordering, and data grouping affect estimation accuracy in waffle charts with a two-value containment structure. Data from 73 participants who completed a 12-condition within-subject study were analyzed using mixed-effects multinomial logistic regression. Visual mark arrangement and color ordering showed no statistically significant effects, whereas data grouping significantly affected accuracy. Participants performed more accurately when the two quantities differed substantially and estimated the smaller contained value more accurately than the larger containing value. These findings indicate that relationships between data magnitudes may affect estimation accuracy more strongly than the tested visual-design variations.

Index terms: Human-centered computing, Visualization, Empirical studies in visualization.

1 INTRODUCTION

Waffle charts represent proportions through a regular grid of equally sized cells and support quantitative estimation through counting or approximate numerosity. Although they are commonly used to communicate part-to-whole relationships, some designs present multiple values within the same grid, creating a containment structure in which one quantity is visually enclosed within another.

This study was motivated by a recruitment-funnel example in *Storytelling with Data* [1], which represents 100 phone screenings, 25 on-site interviews, and 9 job offers as nested regions within a single waffle chart (Figure 1). The extent to which viewers can accurately estimate these nested quantities and perceive their containment relationship without numerical labels remains unclear.

Research on graphical perception indicates that quantitative accuracy depends on both visual encoding and perceptual task. Studies of visual metaphors further suggest that enclosure and containment influence the interpretation of relationships among graphical elements [2]. Prior work on related unit-based displays has shown that the spatial arrangement of filled marks can affect proportion estimation in icon arrays [3]. Building on these findings, this study examines whether visual mark arrangement similarly affects the estimation of two values presented in a containment structure. Data grouping was included to examine whether greater numerical separation improves the discriminability of the two quantities, while color ordering was examined as an exploratory factor to assess whether the relative lightness of the contained and containing regions affects estimation accuracy.

2 STUDY DESIGN

A power analysis with $\alpha = .05$, power = .80, and a medium effect size (Cohen's $f = 0.25$) indicated that at least 63 participants were

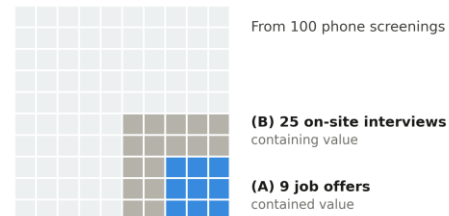


Figure 1: A waffle chart illustrating a recruitment funnel with a containment structure. Reproduced from *Storytelling with Data* [1]

required. Participants were recruited through convenience and snowball sampling; screen size and resolution were not standardized. Of the 80 participants who completed all 12 trials, seven were excluded for out-of-range responses or unequal coverage of the three data groups, leaving 73. No attention checks were used.

Each stimulus consisted of a 10×10 waffle chart containing 100 cells. The study examined three factors. Visual mark arrangement had two levels: horizontal and square. In the horizontal arrangement, filled cells were organized across contiguous rows, whereas the square arrangement formed compact blocks. Color ordering had two levels: light-to-dark and dark-to-light. The same sequential color palette was used across all conditions so that only the ordering of light and dark colors varied.

Data grouping consisted of low-low, low-high, and high-high pairs. The low-value set was $\{4, 9, 16, 25\}$, and the high-value set was $\{36, 49, 64, 81\}$. The values were selected as perfect squares to support compact square configurations without partial rows. Crossing two arrangements, two color orders, and three data groups produced 12 experimental conditions, as shown in Figure 2.

Participants entered numerical estimates for the contained value (A) and the containing value (B). The experiment was implemented using reVISit [4] and deployed online for desktop or laptop use. Condition presentation followed a Latin Square design in which each participant saw 12 stimuli, one from each cell of the $2 \times 2 \times 3$ design, and could not return to previous trials. No practice trials were provided.

The study configuration and stimuli are available at github.com/Dheerissara/Study/tree/main/public/waffle-chart-study.

3 RESULTS AND DISCUSSION

An estimate was classified as correct when it fell within $\pm 10\%$ of the ground-truth value. Sensitivity analyses using $\pm 5\%$, $\pm 10\%$, and $\pm 15\%$ produced a consistent pattern. Joint accuracy was classified as CC (both A and B correct), CN (only A correct), NC (only B correct), or NN (neither value correct).

Mixed-effects multinomial logistic regression was used because the dependent variable consisted of four nominal categories and each participant contributed repeated responses. Participant-level random intercepts accounted for within-participant dependence and

* e-mail: 6780188726@cbs.chula.ac.th

† e-mail: puripant@cbs.chula.ac.th

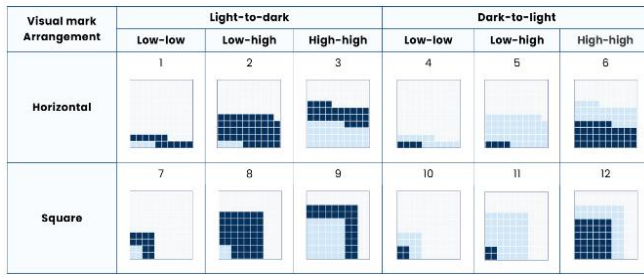


Figure 2: The 12 experimental conditions combining two visual mark arrangements, two color orderings, and three data groups.

individual differences. Generalized estimating equations were used as a robustness check.

Figure 3 presents the proportions of the four joint-accuracy outcomes. CN was the most frequent outcome, accounting for approximately 65% of responses, whereas approximately 28% of responses were correct for both values. Participants therefore estimated the smaller contained value A more accurately than the larger value B.

Horizontal and square arrangements produced similar outcome distributions, as did light-to-dark and dark-to-light color ordering. Neither reached significance in any outcome equation (smallest $p = .073$ for arrangement and $.272$ for color ordering).

Data grouping was the only statistically significant factor. With CC as the reference outcome and high-high as the reference level, low-high pairs reduced the odds of obtaining only A correct ($OR = 0.25$, 95% CI $[0.14, 0.44]$, $p < .001$) and the odds of estimating both values incorrectly ($OR = 0.09$, 95% CI $[0.03, 0.24]$, $p < .001$). Participants were therefore more likely to estimate both quantities correctly when the values differed substantially. The robustness analysis produced the same overall pattern. Interaction effects were not modeled: with one trial per participant per condition, the design is underpowered for such tests.

The results indicate that the numerical relationship between the displayed values affected estimation accuracy more strongly than the tested variations in arrangement and color ordering. Greater separation between A and B may have increased the discriminability of their cell quantities. Although prior work found that spatial arrangement can affect proportion estimation in icon arrays [3], the two arrangements tested in the present containment waffle charts did not produce statistically significant differences.

The predominance of CN suggests that the smaller contained value was easier to estimate than the larger containing value. A possible explanation is that A formed a compact and visually bounded group, whereas estimating B required perceptual integration across a larger surrounding region. Because viewing strategies were not directly measured, this interpretation remains a hypothesis.

For visualization design, the findings suggest that containment waffle charts may communicate part-to-whole structure while making precise estimation of both values difficult. Numerical labels should therefore be considered when accurate reading is required, particularly for the larger value or when the two quantities are similar.

4 CONCLUSION AND FUTURE WORK

Under the tested two-value containment conditions, visual mark arrangement and color ordering did not produce statistically significant differences in estimation accuracy. Data grouping was the only statistically significant factor, with participants performing more accurately when the two quantities differed substantially. The smaller contained value was also estimated more accurately than the larger containing value. These findings indicate that relationships between co-presented data values should be

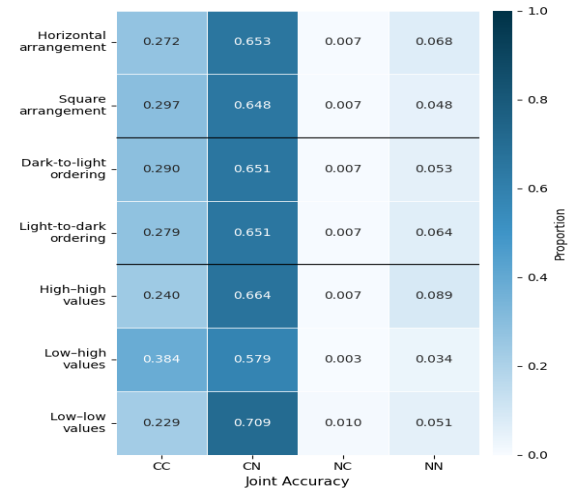


Figure 3: Proportions of joint accuracy outcomes across visual mark arrangement, color ordering, and data grouping conditions.

considered alongside visual-design choices when waffle charts are used for quantitative communication. Numerical labels may therefore be necessary when precise estimation of both values is required.

These findings should be interpreted in light of several limitations. All stimulus values were perfect squares, which allowed the square arrangement to be read by counting one side, so the absence of an arrangement effect applies only to this stimulus set. Accuracy was scored as a binary $\pm 10\%$ outcome, which treats small and large deviations alike and is stricter for low values than for high ones. No practice or attention-check trials were included, so inattentive responses may remain. Demographic information was not collected, which limits the scoping of the results. The experiment examined one fixed color palette and a two-value containment structure, and did not measure comparison tasks, overview tasks, or response time.

Future studies should test values that are not perfect squares, score accuracy as a continuous deviation from the ground truth, and add attention checks and more trials per participant so that interactions can be estimated. How far these results generalize to other unit-based representations, such as icon arrays, remains open. Response-time analysis could provide further evidence regarding task difficulty, while eye-tracking could help determine whether viewers estimate the smaller contained value before integrating the surrounding cells associated with the larger value.

REFERENCES

- [1] C. N. Knaflic, *Storytelling with Data: A Data Visualization Guide for Business Professionals*. Hoboken, NJ, USA: John Wiley & Sons, 2015.
- [2] C. Ziemkiewicz and R. Kosara, "The shaping of information by visual metaphors," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1269–1276, 2008, doi: 10.1109/TVCG.2008.171.
- [3] C. Xiong, A. Sarvghad, D. G. Goldstein, J. M. Hofman, and Ç. Demiralp, "Investigating perceptual biases in icon arrays," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 2022, Art. no. 137, pp. 1–12, doi: 10.1145/3491102.3501874.
- [4] Y. Ding, J. Wilburn, H. Shrestha, A. Ndlovu, K. Gadhave, C. Nobre, A. Lex, and L. Harrison, "reVISit: Supporting scalable evaluation of interactive visualizations," in *Proceedings of the 2023 IEEE Visualization and Visual Analytics (VIS)*, 2023, pp. 31–35, doi: 10.1109/VIS54172.2023.00015.