

Heat3d: A Study on 3D Heatmaps

STAT 992 Final Paper

Tyler Wiederich

May 18, 2025

The display of 3-dimensional (3D) data is often limited by how it can be rendered. These renderings are typically presented as charts on 2-dimensional (2D) computer screens using 2D or 3D chart styles. However, computer renderings do not provide the level of intractability of 3D charts that we experience in a 3D world, something which can be accomplished through the use of 3D printing. In this paper, we describe a study to compare 2D and 3D heatmaps rendered digitally or via 3D printing and assess the findings of 3D printed charts for the use of displaying statistical information.

Introduction

In the 20th century, advancements in technology made data visualizations increasingly more affordable and accessible to a broader population. The primary change in formal data visualizations was from hand-drawn charts to computer-rendered charts (Tukey 1965), yet other technological advances have allowed for data visualizations to enter other mediums. These include the ability to effectively use the 3-dimensional (3D) world around us with the novel use of 3D-printed charts. This newer type of visualization has gained a small amount of traction in recent years as a method of producing tangible charts.

At this time, there are few, if any, studies that evaluate 3D-printed charts for the purpose of displaying statistical information. This may in part be due to the cost of materials and rendering times as compared to charts produced on a computer. A single chart can take up to a day to print, limiting the ability to quickly produce visualizations. Given the nature of 3-dimensional data, viewing a 3D realization of statistical information in a 3D environment is a realistic scenario that could increase the ability for information extraction.

3D-printed visualizations have shown mixed or promising results across many disciplines. Kat-sioloudis and Jones (2018) showed no evidence of a statistical difference in the method of 3D renderings when tasked with drawing a cross-sectional of a dodecahedron. This demonstrated

that spatial awareness between computer-rendered and 3D-printed shapes is not largely different among engineering students. The use of 3D-printed maps for navigation for visually impaired persons showed positive feedback by Holloway et al. (2019), increasing the accessibility of navigation. In the clinical setting, 3D-printed anatomy structures were well accepted along with VR-glasses and 3D computer renderings (Muff et al. 2022). With the rise of 3D-printed visualizations in scientific fields, we turn our attention to their use for statistical graphics.

Literature Review

An identical dataset across multiple chart types does not ensure that the data is perceived in the same way (Cleveland and McGill 1984; Hofmann et al. 2012; Vanderplas, Cook, and Hofmann 2020). One of the main factors to this phenomena is with how data is encoded into the chart. These encodings include, but are not limited to, placement along axes, lengths, areas, volumes, and color scales. For example, bar charts and pie charts are two common visualizations that have long since been the topic of debate (Eells 1926; Croxton and Stryker 1927; Cleveland and McGill 1984). In the case of bar charts and pie charts, encodings are represented by lengths and angles, respectively. Inherently, the comparison of different chart types is a comparison of encodings due to the changes in how data is being displayed.

Cleveland and McGill (1984) noted that estimates involving numerical accuracy may decrease when increasing dimensionality of the encoding, although this was not formally tested in their experiments. The reasoning is possibly due to Stevens' power law, a mathematical formulation of how magnitudes are perceived given different stimuli sources (Stevens and Stevens 1986). The general form of the law is $\psi(I) = kI^\alpha$, where I is the magnitude of a stimulus, $\psi(I)$ is the perceived magnitude, k is a proportionality constant from the unit of the stimulus, and α is the exponent from the type of stimuli used. Studies have estimated that lengths are perceived without bias (i.e., $\alpha = 1$), but areas and volumes tend to have skewed perceptions (Cleveland and McGill 1984). This indicates that lower-dimensional charts might perform better when readers are tasked with extracting numerical estimates from the chart.

There are mixed results in regard to the use of 3D charts, mostly attributing to the purpose of third dimension. When the extra dimension does not convey meaningful information, estimates of accuracy decrease and solution times increase as compared to equivalent 2D charts (Fisher, Dempsey, and Marousky 1997; Zacks et al. 1998; Fischer 2000). The same increase in solution time is seen when the third dimension is utilized for displaying data, but can sometimes produce better error rates than 2D charts (Barfield and Robless 1989; Kraus et al. 2020). Additionally, when given the option of 2D or 3D charts for extracting numerical information, the 2D charts showed increased preference and confidence than their 3D counterparts (Barfield and Robless 1989; Fisher, Dempsey, and Marousky 1997). It is worth noting that all of the studies listed use renderings of 3D charts and not physical 3D charts.

Formal studies involving true 3D charts are limited, and it is unclear if they follow the framework of existing theories on data visualizations. Unlike paper and computer rendered charts, constructing true 3D charts inherently contains many additional factors that could affect perception, such as chart materials, natural lighting, interactivity, and viewing distance. Some of these factors have already been shown to have an effect on computer renderings (Tarr and Kriegman 2001; Wang et al. 2022).

We hypothesize that 3D charts in 3D environments will produce better information extraction than their computer rendered counterparts. Specifically, we will compare 3D-printed charts to digital 2D and 3D renderings. We suspect that this difference will hold across multiple data sets and different magnitudes of stimuli. In this paper, we evaluate the accuracy of numerical estimations on true 3D charts by conducting a factorial experiment that assessed chart types and ratios of pairs of stimuli. We discuss the construction of the stimuli and how we closely matched the charts to compare 2D, 3D-digital, and 3D-printed renderings of heatmap data.

Methods

Our study is designed to evaluate and expand the literature on numerical estimation of 3D charts. In our study, our focus is on 2D and 3D heatmaps, which we carefully construct to ensure that differences are contributed to the dimensionality of the chart. All of our data and methods are publically available for open science and reproducibility at <https://github.com/TWiedRW/ch3-heat3d>. In this section, we discuss process of designing our experiment and participant recruitment.

Stimuli

We denote “stimuli” to represent the magnitude of our chosen values. In a 3D Cartesian space, the X- and Y- axes represent the coordinates of the stimuli, and the Z-axis represents the value for the stimuli. Each X and Y coordinate is represented by a square tile with a 1:1 aspect ratio. For a 2D space, the Z-axis is replaced by a color gradient scale. All stimuli and remaining randomly generated values range between 0 and 100 units. In this experiment, $X = 1, 2, \dots, 10$ and $Y = 1, 2, \dots, 10$.

The design of our experiment made use of the method of constant stimuli, where comparisons between stimuli are with respect to a stimuli that remains the same magnitude (Kingdom and Prins 2016, chap. 3). We set the constant stimuli at 50 units. For stimuli between 50 and 100, we set the maximum stimuli value at 90 and equally partitioned the ratios of stimuli with the constant stimuli, $\frac{50}{90} = 0.556$ to $\frac{50}{50} = 1.0$, resulting in 4 varying stimuli values where 50 is the smallest value. The same ratios obtained with stimuli between 50 and 90 were used to create 4 stimuli varying between 0 and 50, where 50 is the largest value. Additionally, we also included a stimuli pair where both values are 50, resulting in nine total pairs of stimuli. All stimuli values can be found in Figure 1.

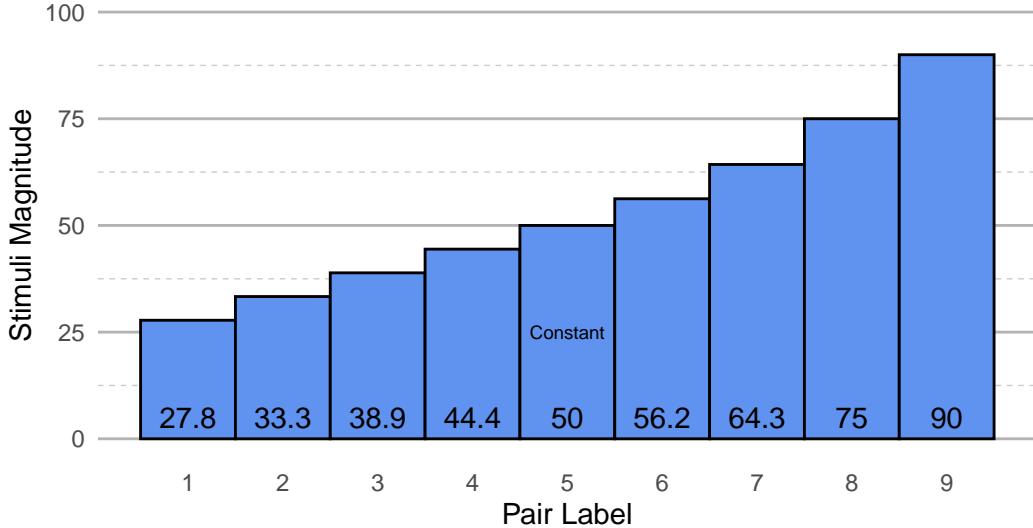


Figure 1: Values for stimuli in the heatmap experiment. All values are paired with the constant stimuli of 50 units, creating nine pairs of stimuli.

To generate non-stimuli random values, we used a mixture distribution of random uniform noise and a mathematical function to populate our coordinate grid. The mathematical functions are scaled between 0 and 100, $g(Z) = 100 \cdot \frac{Z - \min(Z)}{\max(Z) - \min(Z)}$. Two datasets were created for the experiment. The first dataset, called set 1, used the formula for the top half of sphere that is centered within our X and Y coordinate grid, $f_1(X, Y) = \sqrt{7^2 - (X - \bar{X})^2 - (Y - \bar{Y})^2}$, where \bar{X} and \bar{Y} are the averages of the X and Y coordinate ranges. The second dataset is calculated similarly using the formula for the bottom half of sphere, $f_2(X, Y) = \sqrt{7^2 - (X - \bar{X})^2 + (Y - \bar{Y})^2}$. Denoting Z as the random values, $U(0, 100)$ as a random variable drawn from a continuous uniform distribution, and X, Y as heatmap coordinates, our random heatmap data is calculated in Equation 1 with $c = 0.3$. An example of varying c values is presented in Figure 2.

$$Z = c \cdot U(0, 100) + (1 - c) \cdot g(f_i(X, Y)) \quad (1)$$

The placement of stimuli values onto the randomly generated heatmap data was done via simulation to try to make their placement look as natural as possible. Twenty random heatmaps were generated. For each heatmap, the non-constant stimuli was placed onto the coordinate such that the difference between the stimuli and randomly generated value is minimized. The constant stimuli was then placed onto a coordinate with a Manhattan distance of three or four that minimizes the difference between the constant stimuli and randomly generated values, where the Manhattan distance is given by $|X_i - X_j| + |Y_i - Y_j|$ for stimuli i and j . To ensure that stimuli placement is evenly position across the heatmap, the count of stimuli was computed separately across the X and Y axes. For example, in Data Set 1, the X-axis has

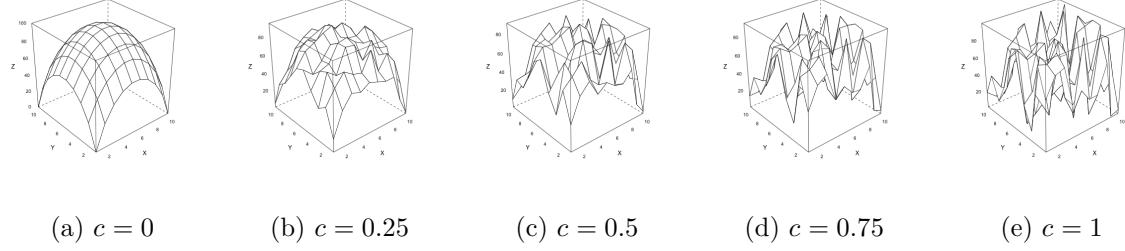


Figure 2: Mixture distribution of Equation 1 using the formula for the top half of a sphere. As c approaches 1, the distribution resembles uniform random noise.

four stimuli in $X = 1$, three stimuli in $X = 2$, and so forth. Chi-squared statistics were calculated for each axis and the heatmap with the smallest average Chi-squared statistic was selected as the final dataset. A visual inspection of this process showed that the stimuli were not clustered in any one area of the chart and that the stimuli look natural with respect to the random mixture distribution.

Charts

Three types of charts were considered for this study: 2D-digital (2dd), 3D-digital (3dd), and 3D-printed (3dp). We constructed these charts so that they are as similar as possible, but inherent difference between dimensionality led to artistic decisions that attempt to focus solely on the dimensionality of the charts. The process of creating the charts is discussed in this section.

The 3D-printed charts were rendered with OpenSCAD (Kintel 2023). To include plot text, a 120mm by 120mm by 10mm base was created with a solid color that was either white or black. Cells of the heatmap measured 10mm by 10mm, resulting in a heatmap that is 100mm by 100mm and is centered on the base. The upper bound of the height of heatmap values is 100mm, where 1-unit in the heatmap data is represented by 1mm of height on the heatmap. Once rendered, the heatmap was saved to 3D Manufacturing Format (3mf) and Standard Triangle Language (stl) files. A variety of solid and gradient filaments were used to print the output files from OpenSCAD. An example of the 3D-printed chart is shown in Figure 3.

To closely match the 3D-digital chart to the 3D-printed chart, multiple stl files were created for each colored component and combined with the RGL package (Murdoch and Adler 2023). The base was rendered with white smoke (#F5F5F5) to slightly contrast with the default white background color (#FFFFFF). Heatmap tiles were rendered with cyan (#74CCFF) and text labels were rendered with black (#000000). Lighting was fixed at 45 degree angles at two opposite corners of the chart. The end result was a near perfect replica of the 3D-printed charts, with the exception of different heatmap tile colors, where an example is given in Figure 4.



(a) Solid filament of Data Set 1

(b) Solid filament of Data Set 2

Figure 3: 3D printed heatmaps.

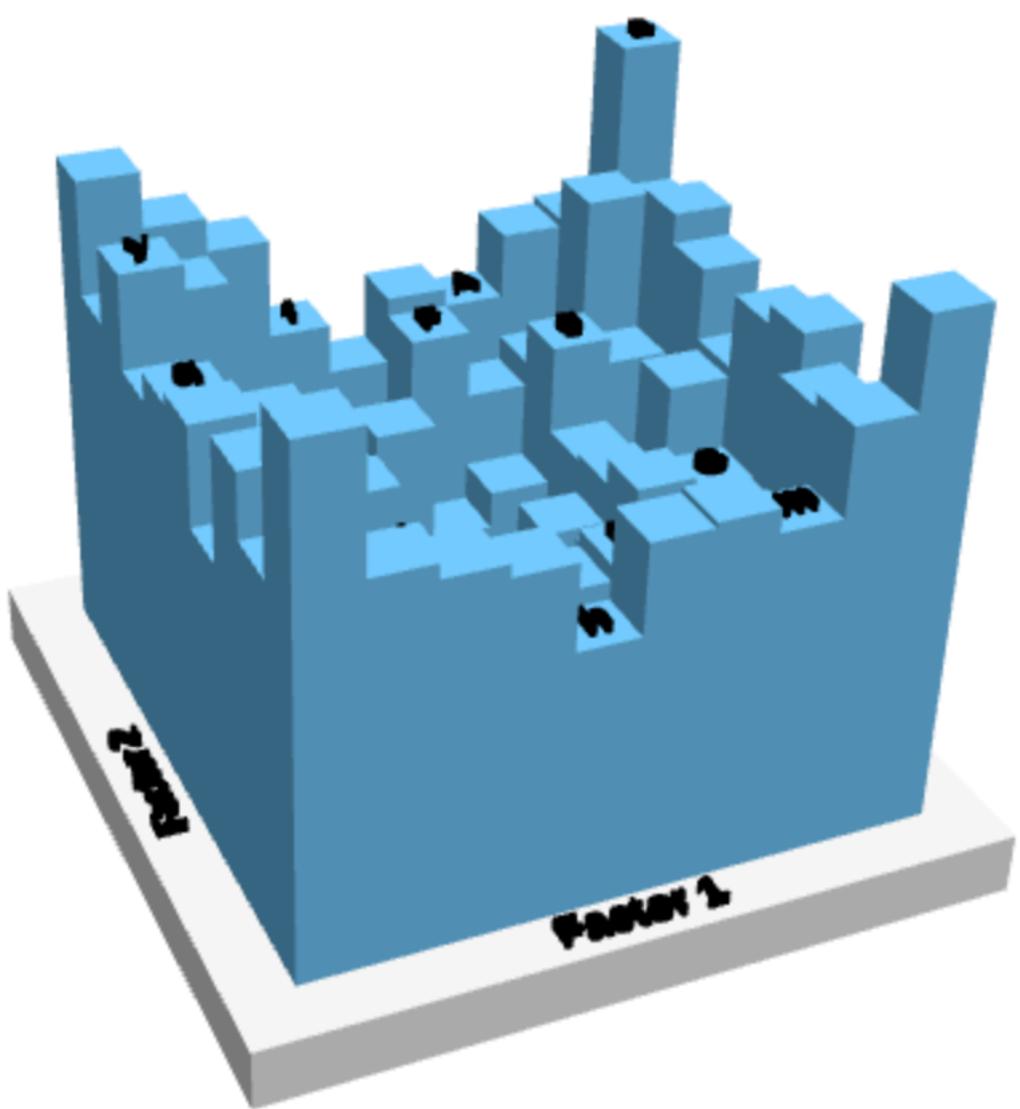


Figure 4: RGL rendering of Data Set 2 for 3dd charts.

Unlike the 3D charts, the 2D charts needed a different encoding to convey heatmap values. The 2D heatmaps were created with `ggplot2` (Wickham 2016) using `geom_tile()`. Fill colors for the cells use a color gradient from Blue Zodiac (#0C2841) to Malibu (#66D9FF), which were selected from a color picker using shadows on our initial 3D-printed charts. The color interpolation was performed with the `scale_fill_gradient()` function from the `ggplot2` package.



Figure 5: Color palette for 2D-digital charts. The colors are interpolated from #0C2841 to #66D9FF, which are the the colors of the lighting conditions for a 3D-printed chart created with cyan filament.

Subject Recruitment

Participants were recruited from all STAT 218: *Introduction to Statistics* sections at University of Nebraska-Lincoln from June 2025 to Spring 2026. The experiment was incorporated into the curriculum as a project that allowed students to get hands-on experience with statistics. Throughout the project, students were also given a series of reflections to evaluate their perspectives on statistics. The first reflection had gathered students' thoughts on the scientific process before participating in the experiment. After the experiment participation, students reflected on what they thought the experiment was measuring. Lastly, students were given a written article and video presentation of the experiment and asked how their view of the experiment changed from their initial perspectives of the experiment's purpose.

To assess whether students completed the experiment or not, the experiment produced a completion code for students to submit to Canvas. The completion code was created by concatenating three words with dashes (e.g., palm-raising-creatively). Words for the completion code came from the movie transcripts of the Star Wars prequel movies (Episodes 1, 2, and 3) as presented by movies.fandom.com. The movie transcripts were tokenized and filtered to remove colorful language¹, stop words², and fictional words³. Instructors were provided an R script to check if their students completed the experiment, but the completion codes were not saved with any identifying information that could link experiment responses to student reflection responses.

¹<https://www.cs.cmu.edu/~biglou/resources/bad-words.txt>

²tm package (Feinerer and Hornik 2024)

³Apple Inc. (n.d.)

Experimental Design

Our experiment was designed with a $3 \times 2 \times 9$ treatment structure. Media type is our main interest, with 2D-digital, 3D-digital, and 3D-printed charts. To ensure that results are not confounded with datasets, we used two datasets to create the heatmaps. A total of 9 stimuli pairs were placed into each dataset. The order of treatments was given so that media and dataset combinations were grouped together randomly in the sequence and stimuli pairs were randomized within the groupings.

Due to practical constraints, stimuli pairs were incompletely blocked. A full factorial design would result in 54 trials per participant, which could lead to a decrease in quality responses (Herzog and Bachman 1981). Therefore, we selected four out of the nine possible stimuli pairs to create incomplete blocks. This resulted in 18 blocks for a balanced incomplete block design. Within each block, media type is fully crossed with dataset. Using the incomplete block structure, participants completed a total of 24 trials, which is more practical than the full factorial design. Blocks were chosen for each participant by using probabilities that are proportional to the inverse counts of used blocks, meaning

We measured three responses – two questions for each trial and one question for each media by dataset combination. For each trial, participants are initially asked which stimuli in a pair is larger or if the stimuli are the same value. Next, they were asked to estimate the magnitude of the smaller stimuli if the larger stimuli represents 100 units, which is a subtractive process (Veit, n.d.). After each grouping of media and dataset, participants were presented with a modal dialog window asking them to rate their confidence on a 5-point Likert scale for their answers of the previous group of charts.

Shiny Application

A Shiny application (Chang et al. 2024) was developed to administer the experiment. The application consisted of five sections: informed consent, demographics, practice, experiment, and wrap-up. The entire application is designed to be completed in approximately 30 minutes.

The Shiny application started with the informed consent screen, allowing participants to select if they are a STAT 218 Student and if they agree to the data collection. Participants had to select a data collection option to continue. After submitting their data collection response, a completion code was generated and saved on the last page of the application. A copy of the informed consent is available in our GitHub repository.

If participants agreed to have their data collected, they were presented with the demographics section. This section asked participants to use drop down menus to specify their age, gender identity, highest education level, and a question about how their participation in the experiment is graded. The last question was a text box and asked participants to specify their favorite movie and/or actor. Demographic information was combined with the application

start time and completion code to create a hash for an anonymized participant identifier using the `rlang` package (Henry and Wickham 2025).

Once a participant completed the demographics page or selected “No” to the data collection question, participants were given a practice page. A modal dialog window was initially shown with instructions, and users could display this window again at any point during the practice. The practice consisted of four questions: two trials from 2dd charts and two trials from 3dd charts. Each practice trial showed the correct solution after the participant submits their trial. After all practice trials were completed, another modal dialog window was displayed to ask participants if they have access to the 3D-printed charts. This was necessary since the experiment was given to potentially online sections of STAT 218 who did not have access to the physical charts.

The experiment page was presented to participants after completion of the practice trials Figure 6. Each page contained two questions – one question for identifying which stimuli is larger and another question for estimating the value of the smaller stimuli if the larger stimuli is 100 units. If participants selected that the stimuli are the same value, then the slider was automatically placed at 100 units. Each trial has the slider initially placed at 50 units. An option was provided for displaying the coordinates of the stimuli pair on a 2D heatmap, but was left unchecked by default. For 3dd charts, the number of interactive clicks was also recorded. A trial could only be submitted if the first question is answered and if the slider was moved at least once.

Trial 1 of 24

Group 1 of 6

Which of the following values is larger?

- F
- f
- The two values are the same.

If the larger value you selected above represents 100 units, how many units does the smaller value represent?

Submit

Select this checkbox if you need help identifying the two values on the chart.

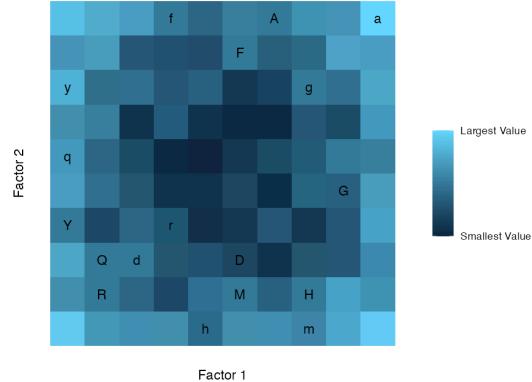


Figure 6: Graphics experiment page. Each trial contained two questions about a pair of stimuli, and a check box that displays the location of the stimuli if needed.

After completing all trials, participants were given the completion code and informed to copy

the code to Canvas since they will not have access to the code after closing out of the experiment.

Method of Analysis

The design of our experiment is a balanced incomplete split-plot with replication, where media type and dataset is the whole plot and stimuli pair is the split plot. A similar design was presented by Mandal, Parsad, and Dash (2020) for a single replicate. For our experiment, the ANOVA table is in Table 1.

Table 1: ANOVA table for analysis of a single replicate of the balanced incomplete split-plot.

Source	Degrees of Freedom
Block (B)	18-1
Media type (M)	3-1
Data Set (D)	2-1
Whole-plot error = $B \times M \times D$	(18-1)(3-1)(2-1)
Stimuli Pair (P)	9-1
$P \times M$	(9-1)(3-1)
$P \times D$	(9-1)(2-1)
$P \times M \times D$	(9-1)(3-1)(2-1)
Split-plot error = $B \times M \times D \times P$	(18-1)(3-1)(2-1)(9-1)

The first question for each trial was to indicate which value within a stimuli pair was larger, or if they were the same value. Since a participant could either answer this question correctly or incorrectly, a Binomial generalized linear mixed model (Stroup et al., n.d.) was fit to the responses.

To measure numerical accuracy of participant responses for the second question in each trial, we used the absolute value of the difference between the participant estimate of the smaller stimuli value and its true value. By using the absolute value of the error, we measure how far off a participant's estimate was rather than the direction of error. For numerical accuracy, we fit a linear mixed model.

$$Y = |\text{Guess} - \text{Actual}|$$

In addition to linear mixed models, we used the bootstrap distribution of mean errors as described by Cleveland and McGill (1984) to provide empirical claims for the effect of dimensionality. A common issue in experiments involving perception is that random noise often covers the effects of interest, so the empirical claims may provide additional insight into the effect of chart dimensionality (XXX).

After each media type and data set grouping, participants were given a question to rate their confidence on their answers. Here, we use descriptive statistics and visualizations to assess confidence across our treatment factors.

Limitations

Our study was designed to evaluate the accuracy of numerical estimation of stimuli pairs across three types of charts. However, numerical accuracy is not the only deciding factor in what makes a good chart design. Other factors include comprehension, recall, and overall preference (Fisher, Dempsey, and Marousky 1997; Vanderplas, Cook, and Hofmann 2020). Given the lack of current research into 3D-printed statistical graphics, we hope to provide some baseline insights into their use as a statistical tool.

We did not account for potential visual illusions in the construction of the heatmap data. A phenomenon seen in statistical graphics studies is that surrounding values may attribute to participant estimates (Zacks et al. 1998; VanderPlas and Hofmann 2015). In our study, the placement of stimuli was so that the stimuli looked naturally positioned within the structure of the data, resembling patterns typically seen in spatial data. However, consistency of neighboring values was not accounted for in this experiment and would be worth having a follow up study to measure the effect of surrounding values on 3D-printed charts.

Lastly, we wanted to construct heatmaps so that they closely resembled each other across dimensionality. The 3dd charts are identical to the 3dp charts, with the exception of heatmap colors. Here, the issue is that 3D heatmaps are not typically constructed by compiling stl files, but rather using packages such like `lattice` package (Sarkar 2008) to display the data. Due to parameter settings and code limitations, it may not be possible to get typical 3D renderings of charts to appear visually identical to 3D-printed counterparts.

Conclusion

Our study provides insight into the novel use of 3D-printed statistical graphics. We hope that these charts may be useful for educational purposes, accessibility of visually impaired persons, or recreational use of statistical communication.

- Apple Inc. n.d. “Web2 Dictionary File.” </usr/share/dict/web2>.
- Barfield, Woodrow, and Robert Robless. 1989. “The Effects of Two- or Three-Dimensional Graphics on the Problem-Solving Performance of Experienced and Novice Decision Makers.” *Behaviour & Information Technology* 8 (5): 369–85. <https://doi.org/10.1080/01449298908914567>.

- Chang, Winston, Joe Cheng, JJ Allaire, Carson Sievert, Barret Schloerke, Yihui Xie, Jeff Allen, Jonathan McPherson, Alan Dipert, and Barbara Borges. 2024. “Shiny: Web Application Framework for r.” <https://shiny.posit.co/>.
- Cleveland, William S., and Robert McGill. 1984. “Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods.” *Journal of the American Statistical Association* 79 (387): 531–54. <https://doi.org/10.1080/01621459.1984.10478080>.
- Croxton, Frederick E., and Roy E. Stryker. 1927. “Bar Charts Versus Circle Diagrams.” *Journal of the American Statistical Association* 22 (160): 473–82. <https://doi.org/10.2307/2276829>.
- Eells, Walter Crosby. 1926. “The Relative Merits of Circles and Bars for Representing Component Parts.” *Journal of the American Statistical Association* 21 (154): 119–32. <https://doi.org/10.1080/01621459.1926.10502165>.
- Feinerer, Ingo, and Kurt Hornik. 2024. “Tm: Text Mining Package.” <https://CRAN.R-project.org/package=tm>.
- Fischer, Martin H. 2000. “Do Irrelevant Depth Cues Affect the Comprehension of Bar Graphs?” *Applied Cognitive Psychology* 14 (2): 151–62. [https://doi.org/10.1002/\(SICI\)1099-0720\(200003/04\)14:2%3C151::AID-ACP629%3E3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1099-0720(200003/04)14:2%3C151::AID-ACP629%3E3.0.CO;2-Z).
- Fisher, Samuel H., John V. Dempsey, and Robert T. Marousky. 1997. “Data Visualization: Preference and Use of Two-Dimensional and Three-Dimensional Graphs.” *Social Science Computer Review* 15 (3): 256–63. <https://doi.org/10.1177/089443939701500303>.
- Henry, Lionel, and Hadley Wickham. 2025. “Rlang: Functions for Base Types and Core r and ‘Tidyverse’ Features.” <https://rlang.r-lib.org>.
- Herzog, A. Regula, and Jerald G. Bachman. 1981. “Effects of Questionnaire Length on Response Quality.” *The Public Opinion Quarterly* 45 (4): 549–59. <https://www.jstor.org/stable/2748903>.
- Hofmann, Heike, Lendie Follett, Mahbubul Majumder, and Dianne Cook. 2012. “Graphical Tests for Power Comparison of Competing Designs.” *IEEE Transactions on Visualization and Computer Graphics* 18 (12): 2441–48. <https://doi.org/10.1109/TVCG.2012.230>.
- Holloway, Leona, Kim Marriott, Matthew Butler, and Samuel Reinders. 2019. “ASSETS ’19: The 21st International ACM SIGACCESS Conference on Computers and Accessibility.” In, 183–95. Pittsburgh PA USA: ACM. <https://doi.org/10.1145/3308561.3353790>.
- Katsioloudis, Petros, and Mildred Jones. 2018. “A Comparative Analysis of Holographic, 3D-Printed, and Computer-Generated Models: Implications for Engineering Technology Students’ Spatial Visualization Ability.” *Journal of Technology Education* 29 (2): 36–53. <https://doi.org/10.21061/jte.v29i2.a.3>.
- Kingdom, Frederick A. A., and Nicolaas Prins. 2016. *Psychophysics: A Practical Introduction*. Second edition. Amsterdam: Elsevier/Academic Press.
- Kintel, Marius. 2023. “OpenSCAD. OpenSCAD.org.” July 17, 2023. <http://openscad.org/documentation.html>.
- Kraus, Matthias, Katrin Angerbauer, Juri Buchmüller, Daniel Schweitzer, Daniel A. Keim, Michael Sedlmair, and Johannes Fuchs. 2020. “CHI ’20: CHI Conference on Human Factors in Computing Systems.” In, 1–14. Honolulu HI USA: ACM. <https://doi.org/10.1145/3313831.3376105>.

[1145/3313831.3376675](https://doi.org/10.1145/3313831.3376675).

- Mandal, B. N., Rajender Parsad, and Sukanta Dash. 2020. “Incomplete Split-Plot Designs: Construction and Analysis.” *Statistics & Probability Letters* 166 (November): 108869. <https://doi.org/10.1016/j.spl.2020.108869>.
- Muff, Julian Louis, Tobias Heye, Florian Markus Thieringer, and Philipp Brantner. 2022. “Clinical Acceptance of Advanced Visualization Methods: A Comparison Study of 3D-Print, Virtual Reality Glasses, and 3D-Display.” *3D Printing in Medicine* 8 (January): 5. <https://doi.org/10.1186/s41205-022-00133-z>.
- Murdoch, Duncan, and Daniel Adler. 2023. *Rgl: 3D Visualization Using OpenGL*.
- Sarkar, Deepayan. 2008. “Lattice: Multivariate Data Visualization with r.” <http://lmdvr.r-forge.r-project.org>.
- Stevens, S. S., and Geraldine Stevens. 1986. *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*. New Brunswick, U.S.A: Transaction Books.
- Stroup, Walter W, George A Milliken, Elizabeth A Claassen, and Russell D Wolfinger. n.d. *SAS for Mixed Models: Introduction and Basic Applications*.
- Tarr, Michael J, and David J Kriegman. 2001. “What Defines a View?” *Vision Research* 41 (15): 1981–2004. [https://doi.org/10.1016/S0042-6989\(01\)00024-4](https://doi.org/10.1016/S0042-6989(01)00024-4).
- Tukey, John W. 1965. “The Technical Tools of Statistics.” *The American Statistician* 19 (2): 23–28. <https://doi.org/10.2307/2682374>.
- Vanderplas, Susan, Dianne Cook, and Heike Hofmann. 2020. “Testing Statistical Charts: What Makes a Good Graph?” *Annual Review of Statistics and Its Application* 7 (1): 61–88. <https://doi.org/10.1146/annurev-statistics-031219-041252>.
- VanderPlas, Susan, and Heike Hofmann. 2015. “Signs of the Sine Illusion—Why We Need to Care.” *Journal of Computational and Graphical Statistics* 24 (4): 1170–90. <https://doi.org/10.1080/10618600.2014.951547>.
- Veit, Clairice T. n.d. “Ratio and Subtractive Processes in Psychophysical Judgment.”
- Wang, Xiyao, Lonni Besançon, Mehdi Ammi, and Tobias Isenberg. 2022. “Understanding Differences Between Combinations of 2D and 3D Input and Output Devices for 3D Data Visualization.” *International Journal of Human-Computer Studies* 163 (July): 102820. <https://doi.org/10.1016/j.ijhcs.2022.102820>.
- Wickham, Hadley. 2016. “Ggplot2: Elegant Graphics for Data Analysis.” <https://ggplot2.tidyverse.org>.
- Zacks, Jeff, Ellen Levy, Barbara Tversky, and Diane J. Schiano. 1998. “Reading Bar Graphs: Effects of Extraneous Depth Cues and Graphical Context.” *Journal of Experimental Psychology: Applied* 4 (2): 119–38. <https://doi.org/10.1037/1076-898X.4.2.119>.