# Evidence for Area as the Primary Visual Cue in Pie Charts

Robert Kosara *


## Abstract

The long-standing assumption of angle as the primary visual cue used to read pie charts has recently been called into question. We conducted a controlled, preregistered study using parallel-projected 3D pie charts. Angle, area, and arc length differ dramatically when projected and change over a large range of values. Modeling these changes and comparing them to study participants' estimates allows us to rank the different visual cues by model fit. Area emerges as the most likely cue used to read pie charts.


## 1 Introduction

Pie charts encode the percentage value to be shown in angle, area, and arc length. Until recently, angle was generally assumed to be the main cue. By asking study participants to estimate values from specially designed pie-like stimuli that isolated the three cues, recent studies have cast doubt on that assumption [11, 20]. Angle by itself turned out to be the least accurate, but there was no clear distinction between area and arc length.

In the study presented in this paper, we take advantage of the distortions introduced by three-dimensional pie charts. Central angle and arc length are distorted by factors of ten and more by the projection (Figures 1 and 2), which can be compared to the responses from study participants. When using parallel projection, a slice's area as a fraction of the entire pie is constant and independent of the view angle (see supplemental material for details), which separates area in particular from the other two cues.

Figure 2 shows the distortion of the projected angle when rotating a 30° slice around a pie chart. Compared to a 2D chart, at a 15° view angle, the slice's central angle increases up to three times its base value, and down to about one quarter (for a total range of almost 12x), arc length increases to four times its base value and down to about one tenth (for a total range of almost 40x).

We are not interested in 3D pie charts themselves or their particular properties here (though we do have some results). Rather, we use them as a sort of model organism that allows us to study a property we want to learn about; similar to the fruit fly in biology.

We model possible responses to stimuli as follows. Detailed formulas are included in the supplemental materials.

1. *Area Model.* Since area of the slice as a fraction of the whole is invariant to the projection, this is the same as using the original angle or the represented percentage to determine the estimate.
2. *Projected Angle Model.* This model predicts that study participants will read the value directly from the projected angle they see. Since angle is strongly distorted by the projection, the difference between this and the original angle model can be detected.
3. *Projected Arc Length Model.* Similar to angle, arc length is distorted by the projection. This model predicts that viewers read pie charts using the arc on the outside of the slice, and would thus report values proportional to it.

*Tableau Research, E-Mail: rkosara@tableau.com

Models with fewer inputs are considered better than ones with more, which is also directly expressed in criteria for model selection like the Akaike Information Criterion (AIC) we use below. Of the three models above, the area model depends only on a single input, the represented value. The other two have to take into account the view angle and rotation of the chart around the center in addition to the value, so both depend on three input parameters.

## 2 Related Work

Pie charts are among the classic visualization techniques [21]. William Playfair included early versions in his 1801 Statistical Breviary [15] and added them as an illustration to a statistical account of the United States he translated [8].

Below, we briefly survey the limited literature on the perceptual basics of pie charts, including 3D pie charts.

### 2.1 Pie Chart Perception and Effectiveness

Visualization books generally state angle as the method by which we read pie charts, often without giving a reference. This dates back to Brinton in 1914 [3], and continues with Bertin [1] and much more recent books such as those by Robbins [17], Munzner [13], etc. Cleveland and McGill [6] also equate their pie chart stimulus with angle perception without questioning it, as do Simkin and Hastie [19], and others.

The basis for the angle assumption appears to be Eells' 1926 paper [9]. Eells conducted a study comparing bar to pie charts, and asked his participants how they had read the pie charts. Just over half of them stated angle as the mechanism, with the other half being split between area and arc length. Self-reported strategies can be unreliable, however, and would at least require additional studies to corroborate before turning into the canonical assumption.

Skau and Kosara conducted a series of studies that cast doubt on the angle assumption. By deconstructing pie charts into individual visual cues, they were able to collect evidence that angle was the least likely visual cue [20]. They also hypothesized that a pie chart with a larger slice would lead to overestimation because its arc length and area were larger, and found that to be the case [11]. They were unable to differentiate between arc length and area, however.

The classic Cleveland and McGill paper [6] found pie charts to be worse than pure bar charts, but similar to some bar chart configurations (such as stacked bars). Earlier studies in the 1920s had found different results, though, in particular for part-whole judgments. Eells [9] found pie charts to be superior to bars, and von Huhn [22] was unable to draw a clear conclusion in his studies. Croxton et al. [7] compared a number of specific values (such as 25%, 33%, etc.) and found mixed results: in some cases stacked bars did better, in others the pie chart.

### 2.2 Pie Charts in Three Dimensions

Just as with regular pie charts, the literature is mixed on their 3D brethren. Carswell et al. [5] studied 2D and 3D bar, line, and pie charts. Only 3D line charts were found to have significantly higher error than their 2D counterparts. Siegrist [18] later found a significant effect when comparing 2D to 3D pie charts.

Reading values from 3D pie charts requires the understanding of depth. A study by Lind et al. [12] showed significant error when judging the shape of objects seen under a view angle of about 15°, the steepest angle included in the studies reported in this paper. The

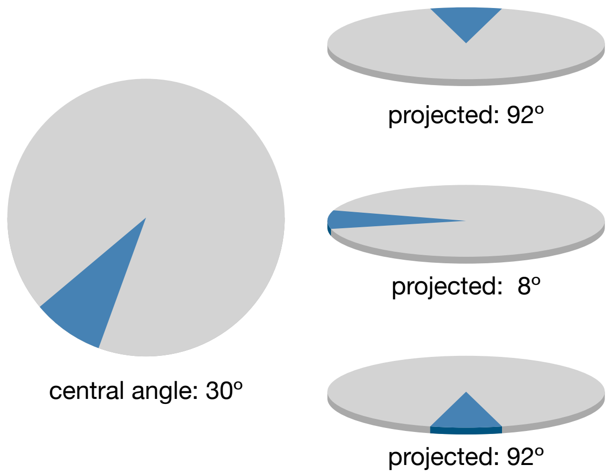


Figure 1: An original *30°* central angle projected at a view angle of *15°* ranges from projected angle of *8°* to *92°*, depending on rotation around the chart. Its fraction of the chart area is constant, however.

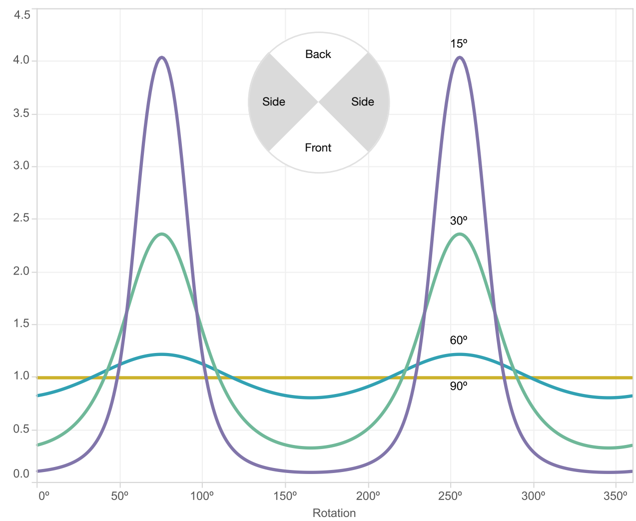


Figure 2: Distortion of arc length as a multiple of the base value for a 30° slice when rotated around a 3D pie chart seen at different view angles. Each line represents a view angle (90° is the 2D case).

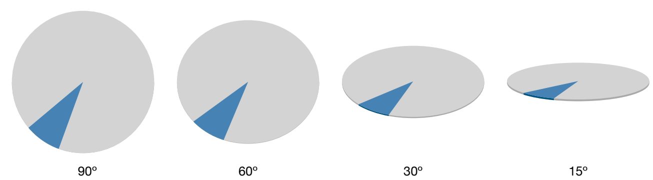


Figure 3: The four view angles used in the study, shown for a body height of 10 pixels.

perception of slant has been shown to be unreliable when respondents had to provide a numerical estimate, but quite accurate when matching a physical replica of the angle [16].

## 3 Study

Given the model predictions, we conducted a controlled online study to determine which model would fit participants' responses.

### 3.1 Materials

Participants were shown pie charts drawn at a width of 600 pixels, their height depended on the view angle. The following factors were varied (see also Figure 3):

- 4 view angles: 90° (2D), 60°, 30°, and 15°
- 3 body heights: 0, 10, and 50 pixels
- 3 value ranges: $< 33\%$, $33\% - 66\%$, and $> 66\%$
- 3 rotations around the center of the pie

This would have yielded $4 \cdot 3 \cdot 3 \cdot 3 = 108$ combinations. However, since body height does not make a difference for the 2D condition, we eliminated height variations for the 90° case and thus reduced the number to 90 ($3 \cdot 3 = 9$ for the 2D condition plus $3 \cdot 3 \cdot 3 \cdot 3 = 81$ for the remaining three view angles).

We pre-generated a set of 90 pseudo-random numbers from the range $[3;97]$, and bucketed them into the three value ranges. During the study, a third of the numbers were picked from each range to ensure equal representation across values. The stimuli were presented in random order.

### 3.2 Procedure

Each step showed a single stimulus and presented a 2D reference chart with a handle that allowed the size of the blue slice to be changed (Figure 4). Participants were asked to mirror the value they were seeing in the stimulus on the reference chart (this *direct report* is common in vision science [2]). The goal was to eliminate the rounding effects seen in earlier studies and avoid guesses for the wrong part of the chart.

The input chart on the right was drawn in the same colors and size as the stimulus. The blue slice started at the 12 o'clock position and was always shown with an initial value of 50% to avoid biasing towards low or high values. There was no numerical display of the chosen percentage. Participants had to change the value before they could advance to the next question. The study consisted of a total of 90 questions, we presented a pause screen after 30 and 60 questions.

### 3.3 Participants

We recruited 80 participants (43 women, 37 men) on Prolific[1], an online platform focused on running studies. A recent study showed its results to be at least comparable to Mechanical Turk [14]. We used Prolific's filters to select only participants with (self-reported) normal or corrected-to-normal vision.

Participants were paid \$2.50 for participation. Study duration was just under 14.5 minutes on average (median 12:53), resulting in an average hourly pay of \$10.42 (median \$11.64).

### 3.4 Pilot Study

We ran a pilot study using the same stimuli but a numerical input instead of the reference chart. We only report the results of this pilot in Table 1, which shows that while it has higher error due to rounding effects from the numerical input, its results are consistent with the main study's.

### 3.5 Preregistration

We decided to preregister this study[2]. Preregistration is becoming a common practice in vision science, perceptual psychology, and other fields to counter the problems of p-hacking and to increase replicability [10]. We lay out the key points of the preregistered

[1] `https://www.prolific.co`
[2] `https://osf.io/7y842/`, also includes study data and code

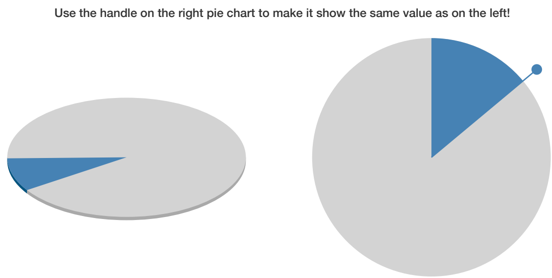


Figure 4: User interface used for the main study. The left plot is the stimulus, the plot on the right has a handle that allows participants to match the value being shown.

analysis below. Section 4.1 reports strictly on the results based on it, with the following sections adding further analyses.

The preregistration is based on the pilot and lists the factors described above, as well as the estimated percentage as the main dependent variable. In the analysis, we give the Akaike Information Criterion (AIC [4], Section 4.1) as the means of distinguishing the models. We had initially included a number of regression models in our analysis of the pilot, but realized later that other measures were more informative. We therefore report AIC below and add the other criteria in the further analysis.

## 4 Results

To determine which model best fits participants' responses, we calculate the difference between the response and the predicted value by each model and trial based on the stimulus' parameters: displayed percentage, view angle, and rotation. Below we use the terms *residual* and *error* interchangeably for this difference.

To compare between the models, we compute two metrics across all trials for each model. The mean average error, MAE, is the mean of the residuals (or errors). The root mean square error, RMSE, is the square root of the mean of all residuals, squared. RMSE is more commonly used in modeling, but it overly weighs outliers and is not as directly interpretable as MAE.

Table 1 shows the values for both MAE and RMSE for all models. Since a better model leads to smaller residuals, smaller values in both measures indicate better models. The values differ between RMSE and MAE (which indicates the presence of outliers), but the relationships between the different models in each are similar. And in each case, the area model has the lowest scores.

### 4.1 Model Evaluation from Preregistered Analysis

Following the analysis laid out in the preregistration above, we find that according to the AIC, the area model has the best fit, followed by the projected angle and arc length models (Table 1).

We calculate AIC as $AIC = 2k + n\ln(RSS)$, with $k$ being the number of parameters in the model, $n$ the number of data points, and RSS being the residual sum of squares (i.e., the sum of the squares of the residuals). RSS is closely related to RMSE, $RMSE = \sqrt{\frac{1}{n}RSS}$.

Since we use the same dataset for all three models, $n$ is identical for all three AIC calculations. The number of parameters $k$ also has a minor influence because of the large value for RSS. The AIC values are therefore mostly determined by the residuals, which we calculate from the uncorrected responses. They agree with the other results (based on the corrected values) that area is the best-fitting model, with angle second and arc length third.

### 4.2 Further Model Evaluation

Building on the preregistered analysis, we add additional steps here for two reasons: the unexpected number of answers for the wrong slice, and the realization that the AIC is not nearly as informative as MAE and RMSE. The key results are not impacted by this, however.

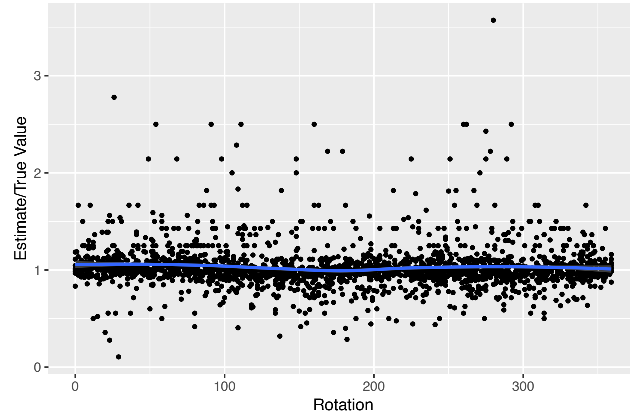


Figure 5: Responses as multiples of the represented value by rotation angle for the 15° condition. The blue line shows a LOESS smoothing (95% CI is invisible at this size). Compared to Figure 2, none of the expected distortions from projected angle or arc length are visible.

| | Pilot Study | | Main Study | | |
|---|---|---|---|---|---|
| **Model** | **MAE** | **RMSE** | **MAE** | **RMSE** | **AIC** |
| Area | 3.3 | 4.7 | 2.2 | 3.2 | 96377 |
| Angle | 5.3 | 7.2 | 4.4 | 6.2 | 98329 |
| Arc Length | 6.8 | 9.1 | 5.9 | 8.4 | 100000 |

Table 1: Comparing the fit of the three predictive models to participants' responses using mean absolute error (MAE), root mean square error (RMSE), as well as Akaike Information Criterion (AIC) for the pilot and the main study. Both studies show area as the model with the best fit (lowest values). AIC is reported for uncorrected values (following the preregistered protocol), others on corrected.

We found that despite the direct report input method, about 2.8% (204 out of 7200) of responses appeared to be for the wrong slice. This was due both to a few individuals responding for the wrong slice almost half the time, and more people having responses for the wrong slice. We corrected estimates that apparently were made for the wrong slice, which we determined by comparing the represented value to the estimate and 100-estimate (same as in earlier studies [11, 20]). Since there are only two slices, answering for the wrong slice means crossing the very obvious 50% mark, which appears unlikely.

We then calculated MAE and RMSE values based on that data (Table 1). All three indicators (MAE, RMSE, and the AIC based on the uncorrected values) agree on the ordering of the models in the pilot and the main study. As Figure 5 shows for the (most extreme) 15° condition, we did not observe the distortion expected from either the angle or arc length models based on rotation of the slice.

Since MAE is a simple means of residuals (as opposed to RMSE, which is the root of the sum of squared errors), it is easier to interpret. Compared to area, MAE is roughly twice as high for the angle model, and about 2.7 times as high for the arc length model.

### 4.3 Additional Analysis: Impact of Perceptual Factors

In contrast to the analysis above, the term *error* in the following refers to the difference between participants' estimates and the true value, not the model prediction. We consider two types of error: signed error is the difference between the estimate and the true value, absolute error is the absolute value of signed error. Absolute error is a measure of accuracy, signed error indicates bias.

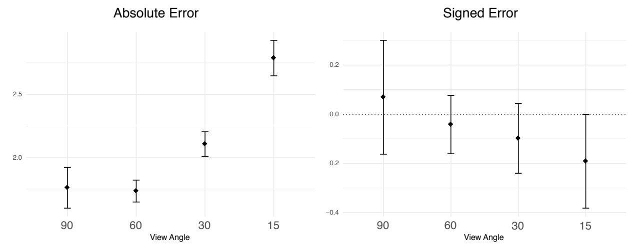


Figure 6: Absolute and signed error by view angle. Absolute error is much higher for the 30° and 15° conditions, but does not differ from the 90° (2D) condition when the angle is 60°.

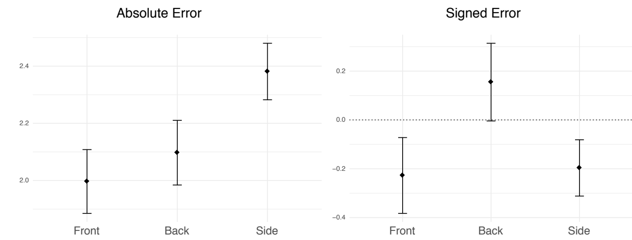


Figure 7: Absolute and signed error by slice orientation. Absolute error for slices on the sides is higher than for ones pointing forward or back. Values for both front- and side-facing slices are underestimated.

### View Angle

The obvious factor in 3D pie charts is the view angle, and we expected it to have a clear effect on people's accuracy. Figure 6 shows 95% confidence intervals (CIs) for both signed and absolute error by view angle. In terms of absolute error, the 90° and 60° conditions do not differ, but the 30° and 15° ones show significantly larger error.

Signed error CIs overlap zero for all conditions except 15°, meaning that there is no consistent bias. While there appears to be a downwards trend towards shallower view angles, the confidence interval for 15° almost touches the zero line, so we do not consider this conclusive evidence for underestimation.

An ANOVA of absolute error confirms the significant difference between conditions ($F(3,2269) = 67.66, p \ll 0.001$), while it shows no effect for signed error ($p = 0.31$).

### Body Height

We did expect a visible body on a projected pie chart to provide additional depth cues that would help people read the values more accurately. That does not appear to be the case, however. Neither absolute nor signed error show significant differences ($p \approx 0.5$).

### Orientation

We expected the orientation of a slice to impact accuracy because it is the second major factor in a slice's distortion. For our analysis, we classified rotation into three bins (*Front*, *Back*, and *Side*) along 45° angles from the axes (see Figure 2, inset).

Absolute error is similar for front- and back-oriented slices, but significantly higher for slices pointing to the side. Signed error shows a negative bias for slices pointing to the front and side (Figure 7) and indicates that there might be a positive bias for slices pointing to the back (the CI includes zero, however). The negative bias for side-pointing slices is expected, but the one for front-facing ones is the opposite of our expectations (wider angle in the front).

An ANOVA confirms what the CIs show: absolute error differs significantly between orientations ($F(2,2270) = 13.49, p \ll 0.001$), as does signed error ($F(2,2270) = 8.303, p < 0.001$).

## 5 Discussion

Our study shows area as the best-fitting model to people's responses, with angle second and arc length last (Table 1 and Figure 5). Given that the earlier studies found angle by itself to show the most error and arc length and area to be equal, this points to area as the most likely cue. Our results clearly contradict the idea that 3D pie charts can be read purely by the projected angle or arc length.

Design parameters such as tilt have an effect, with the 30° and 15° conditions showing significantly more error than 2D or 60°. Subjectively, the 60° condition does not appear to be compressed, and thus might just be read as a 2D pie chart. The more extreme conditions are visibly distorted, however.

The error associated with slices pointing to the side is expected, but it is interesting to note the difference in the direction of the error for slices pointing towards and away from the viewer. These slices are identical in size, the only difference is in people's attempt to recover the two-dimensional shape from the (pseudo-)3D image. They appear to expect slices in the front to be distorted by perspective and correct for that by underreporting their size, whereas the opposite appears to be happening for slices pointing to the back.

One possible explanation for wrong-slice responses is that users sometimes match the part of the pie that is oriented the same way as the blue reference slice in the input chart (which always started at the 12 o'clock position and grew clockwise). If the gray part of the pie starts near the top of the chart, it seems plausible to consider that closer to the input chart.

### 5.1 Limitations

While we assume that the good fit of the area model means that people use area to estimate the percentage, there is an alternative explanation: they might be re-projecting the distorted shape into 2D in their heads and then use any one of the visual cues to read the value. This is consistent with the area model, since it is also the model that simply uses the original value to predict the estimate.

However, the amount of over- and underestimation depending on slice orientation (Figure 7) point to area: orientation should not impact reading after back-projection. The observed bias seems more likely to be the result of an over-correction after an estimate that does not take the projection angle into account. It is also well established that people's depth perception of shape is very poor [12], thus making it more likely and feasible to estimate area rather than perform a rather complex shape and size estimation on a tilted plane. The lack of an effect from body height supports this, since additional depth cues would aid the back-projection.

The charts used in the studies presented in this paper consist of only two slices: the one indicating the value we ask about and the rest of the chart. Since we are interested in the visual cue that people use to read the chart, and not pie chart effectiveness or the impact of multiple slices, we do not believe this to be an issue.

We chose parallel rather than perspective projection to render the charts in 3D to separate the predicted effects of the different models, which would have been much more complex using perspective projection (and we were not interested in 3D pie charts per se).

## 6 Conclusions

Our study points to area as the visual cue used to read pie charts. It is consistent with previous work preferring angle over arc length and area. We believe that modeling responses and comparing to study results is an interesting approach to study design that could be more commonly employed in visualization studies.

While this study suggests that the charts are read by area, it is not conclusive. In particular, the possibility of pie chart users re-projecting the chart to read them cannot be ruled out. Further experiments are therefore needed to zero in on the exact mechanism by which this common chart type is read.

## REFERENCES


[1] J. Bertin. *Semiology of Graphics*. University of Wisconsin Press, 1983.

[2] T. F. Brady, T. Konkle, and G. A. Alvarez. A Review of Visual Memory Capacity: Beyond Individual Items and Toward Structured Representations. *Journal of Vision*, 11(5):1–34, 2011.

[3] W. C. Brinton. *Graphic Methods for Presenting Facts*. The Engineering Magazine Company, New York, 1914.

[4] K. P. Burnham and D. R. Anderson. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer, 2nd ed., 2002.

[5] C. M. Carswell, S. Frankenberger, and D. Bernhard. Graphing in Depth: Perspectives on the Use of Three-Dimensional Graphs to Represent Lower-Dimensional Data. *Behaviour & Information Technology*, 10(6):459–474, Nov. 1991.

[6] W. S. Cleveland and R. McGill. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[7] F. E. Croxton and R. E. Stryker. Bar Charts Versus Circle Diagrams. *Journal of the American Statistical Association*, 22(160):437–482, 1927.

[8] D. F. Donnant and W. Playfair. *Statistical Account of the United States of America*. Translated From the French by W. Playfair: with an Addition on the Trade to America. 1805.

[9] W. C. Eells. The Relative Merits of Circles and Bars for Representing Component Parts. *Journal of the American Statistical Association*, 21(154):119–132, 1926.

[10] A. Gelman and E. Loken. The Statistical Crisis in Science. *American Scientist*, 102(6):460–465, 2014.

[11] R. Kosara and D. Skau. Judgment Error in Pie Chart Variations. In *Short Paper Proceedings of the Eurographics/IEEE VGTC Symposium on Visualization EuroVis*, pp. 91–95. The Eurographics Association, 2016.

[12] M. Lind, C. Forsell, and G. P. Bingham. The Illusion of Perceived Metric 3D Structure. In *Proceedings Information Visualization*, pp. 51–56, 2002.

[13] T. Munzner. *Visualization Analysis and Design*. A K Peters, 2014.

[14] E. Peer, L. Brandimarte, S. Samat, and A. Acquistic. Beyond the Turk: Alternative Platforms for Crowdsourcing Behavioral Research. *Journal of Experimental Social Psychology*, 70:153–163, 2017.

[15] W. Playfair. *Statistical Breviary*. Shewing, on a Principle Entirely New, the Resources of Every State and Kingdom in Europe. Wallis, London, 1801.

[16] D. R. Proffitt, M. Bhalla, R. Gossweiler, and J. Midgett. Perceiving Geographical Slant. *Psychonomic Bulletin & Review*, 2(4):409–428, 1995.

[17] N. B. Robbins. *Creating More Effective Graphs*. Chart House, 2013.

[18] M. Siegrist. The Use or Misuse of Three-Dimensional Graphs to Represent Lower-Dimensional Data. *Behaviour & Information Technology*, 15(2):96–100, Nov. 2010.

[19] D. Simkin and R. Hastie. An Information-Processing Analysis of Graph Perception. 82(398):454–465, 1987.

[20] D. Skau and R. Kosara. Arcs, Angles, or Areas: Individual Data Encodings in Pie and Donut Charts. *Computer Graphics Forum*, 35(3):121–130, 2016.

[21] I. Spence. No Humble Pie: The Origins and Usage of a Statistical Chart. *Journal of Educational and Behavioral Statistics*, 30(4):353–368, 2005.

[22] R. von Huhn. Further Studies in the Graphic Use of Circles and Bars: A Discussion of the Eell's Experiment. *Journal of the American Statistical Association*, 22(157):31–39, Mar. 1927.