Bars, Pies, Doughnuts & Tables – Visualization of Proportions

Harri Siirtola

Visual Interaction Research Group (VIRG)
Tampere Unit for Human-Computer Interaction (TAUCHI)
School of Information Sciences, University of Tampere
FIN-33014 University of Tampere, Finland
harri.siirtola@sis.uta.fi

Visualization of proportions is one of the most common visualization types encountered in the media. Stacked bar charts, doughnut and pie charts – the most common visualizations of proportions – have all keen advocates and critics, and there are conflicting results about their performance. This study presents an experiment with a basic but ecologically valid task to evaluate the performance of these techniques. The result shows that the stacked bar chart is superior to doughnut and pie charts in task performance. However, 75% of the participants regarded pie charts as the most pleasing or second-pleasing to use, and almost half of the participants (44%) perceived the pie or doughnut chart to be the fastest visualization to understand. Only six participants recognized the bar chart as the fastest technique, and half of them still preferred the pie chart over them. This subjective preference at the expense of performance may explain why pie charts are so widely used in spite of being criticized.

Keywords: information visualization, proportions, user studies

1. INTRODUCTION

One of the most frequent visualization tasks is to show how the parts of something relate to the whole, or the visualization of proportions. Quintessential ‘wholes’ from the business context are budgets, expenses, profits, sales, and market volumes. In addition, newspapers, newscasts, and other media visualize proportions of things from all aspects of life. It is often crucial to understand what is the composition of something.

The most common visualization techniques in the media for proportions are *pie charts* (Fig. 1), *doughnut charts* (Fig. 2), and *stacked bar charts* (Fig. 3). However, the relative merits and demerits of these techniques have raised considerable debate for over a century now. There are conflicting experimental results and loudly communicated opinions from the practitioners:

- Pie charts are categorically bad: ‘*The only thing worse than a pie chart is several of them*’ (Tufte 1983, p. 178), ‘*Save pies for the dessert*’ (Few 2007), ‘*Death to pie charts*’ (Nussbaumer 2011).

- Pie charts are good: ‘*Why Tufte is Flat-Out Wrong about Pie Charts*’ (Gabrielle 2013), ‘*In Defense of Pie Charts*’ (Kosara 2011), Spence and Lewandowsky (1991), Peck et al. (2013).
- Pie charts perform better than stacked bar charts (Eells 1926).
- Stacked bar charts perform better than pie charts (Cleveland and McGill 1984).
- We should really use tables instead of graphics for a small number of proportions (Ehrenberg 1977).

This paper compares how these techniques perform in one particular task: *perceiving the relative order of the parts of some whole*. The task is not about quantitative aspects of the data, nor is it purely a qualitative task. The performance of a numeric table is used as a baseline, although it is interesting to know what is the cost of showing the exact data which is not needed to solve the task.

One common misunderstanding in the pie chart vs. bar chart controversy is the comparison of pie charts to regular bar charts. The point is that a regular, side-by-side bar chart is missing the ‘whole’, the visual

sum of all values – we need to use a *stacked bar chart* to make a fair comparison.

In the earlier studies, the choice of task has proven to be a problem. It is not easy to choose a task that does not favor some of the techniques. In this experiment, we asked the participants to simply name the parts of the visualization in descending order of their proportion. This is an elementary task when the proportions of a whole are visualized. Another important aspect is how to define ‘performance’. In this study, the performance is the combination of a classic triplet: execution time, error rate, and subjective satisfaction.

The rest of this section gives a brief description and history of the visualization methods under study. The following sections describe the experiment, results and conclusion.

1.1. Pie chart

The pie chart (Fig. 1) was invented by William Playfair and appeared first in 1801 in his book (Playfair 1801, insert between pages 12 and 13). While it is generally agreed that the pie chart is an interesting method to display parts of a whole – if properly used – it is also the most controversial visualization of proportions.

Figure 1: Pie chart: the proportions as segments.

Eells (1926) did the first comparisons between pie charts and stacked (‘segmented’) bar charts where the experimental task was to estimate the percentage of the whole represented by an element, as a number. He concluded that the pie charts could be read as rapidly as stacked bar charts, and that the accuracy of the pie charts was better. Von Huhn (1927) criticized Eells’ experimental task, because neither of the visualizations are used for exact data, but for comparisons. In addition, Huhn suspected that the missing scales and labels from the stimulus ruined the ecological validity and distorted the results. Despite Huhn’s criticism, basically the same experimental setup was later repeated in a number of studies.

Spence and Lewandowsky (1991) did a set of experiments and concluded that pie charts compare favorably to bar charts when tasks other than direct magnitude estimation are required. Their tasks included simple comparisons like ‘which is larger,

A or B?’, but also visual computation tasks like ‘which is larger, $A+B$ or $C+D$?’. They also give a comprehensive review of the earlier research on pie charts.

A recent study by Peck et al. (2013) used near-infrared spectroscopy (fNIRS) to measure impact of a visual design on the brain. They compared pie and bar charts and concluded that there is no universal difference between the two. However, their experimental setup has sparked a vivid discussion (Few 2013).

1.2. Doughnut chart

The doughnut chart (Fig. 2) has been called ‘the poor cousin of the pie chart’. A doughnut chart is created simply by making a hole into a pie chart. Doughnut charts are used extensively in the media.

Figure 2: Doughnut chart: proportions as segment size.

However, removing the center part of a pie chart does have benefits, as it allows to use the space for additional information, such as a label or image. In addition, it is not clear how doughnut charts are read – pie charts are mostly based on the angles, but it is not clear if doughnut charts are based on the estimation of angles or areas.

1.3. Stacked bar chart

It is believed that Nicole Oresme (1320–1382), the famous philosopher of the Middle Ages, invented bar charts and used them in his book *On the Latitude of Form* (Barrow 2011, pp. 352–353). Stacking bars on top of each other (vertically or horizontally) creates a stacked bar chart that shows the whole, the parts, and their relationships (Fig. 3).

Figure 3: Stacked bar chart: proportions as bar length.

In a seminal paper by Cleveland and McGill (1984) a set of experiments was made to evaluate the elementary perceptual tasks. The most accurate perceptual task to extract quantitative information from a visualization was ‘position along a common scale’, followed by ‘position along non-aligned scales’. This suggests that bar charts and stacked bar charts should be efficient in performance comparison.

2. METHOD

2.1. Participants

Sixteen students volunteered to participate in the experiment in partial fulfillment of the introductory course in *Interactive Technology*. Their mean age was 29 years, ranging from 19 to 47. Only one of the participants had taken more than one course in statistics, two had taken one, and the rest had taken none. Seven of the participants had normal vision, and the rest had corrected to normal vision. There was an equal number of women and men in the participants.

2.2. Apparatus

The experimental setup was implemented as a Java program run on a Macintosh computer. The program displayed a stimulus, recorded the participants' input and the time spent on a task (in milliseconds) into a text file. Fig. 4 shows the user interface. The input is a string of letters from the beginning of the alphabet, and the program accepts only those letters that appear in the stimulus image. The answer button (on the bottom right, Fig. 4) activates when all the letters appearing in the stimulus have been inserted into the answer field. The program requires another button press before proceeding to the next task, thus allowing the participant to rest between tasks if so desired.

Figure 4: User interface of the GraphTest application.

The stimuli was created with a statistical system R script (R Development Core Team 2014). The visualizations are as simple as possible – no colors, no scales, plain data labels, and all pixels are data-bearing. In addition to the bar, pie and stacked bar chart condition the table of numbers was used as a baseline condition. The numbers were not percentages, but the same random data that was used to generate the graphic stimuli, shown with two decimal places.

2.3. Procedure

Participants first completed a questionnaire for demographics and background information. Next, each participant was given the same instructions (read aloud from a paper), and familiarized with the four visualization types. The functionality of the *GraphTest* application was explained, and the

participants did eight training tasks with it. The participants were encouraged to ask if there was anything unclear in the *GraphTest* application, the visualizations, or the test itself.

The test task was to list the elements in the visualization in a descending order. E.g. in Fig. 1 the correct answer is *ADBCE*. Participants were instructed to perform the tasks as quickly as possible, but not at the expense of the correct answer.

Each participant did the same set of tasks with all four visualizations. Both the task order and the order of visualization types were counterbalanced with Latin square designs. The tasks were created by varying the number of elements in the visualization (between 4 and 7) and the minimum difference between elements (from 12% to 24%, in 4% increments). With these sixteen tasks and four visualization types each participant performed 64 tasks altogether.

After the test tasks the participants filled in a questionnaire about their subjective opinions on the four visualizations and their performance, and were interviewed for additional comments.

The *GraphTest* application recorded the participants' answers and computed the string distance to the correct answer. The distance metric is the number of adjacent letter swaps required to transform the answer to the correct one. Fig. 5 shows all the possible adjacent letter swaps for the string *ABCD*. If the participant responds *BCAD* and the correct answer is *ABCD*, then the distance is 2.

Figure 5: Distance metric: the number of adjacent letter swaps.

2.4. Design

The experiment was a 4×16 within-subjects factorial design. The independent variables were:

- **type** : The type of the visualization (pie, doughnut, bar, table).

- **task** : Sixteen combinations of element count (4, 5, 6, 7) and the minimum difference between elements (12%, 16%, 20%, 24%).

The dependent variables were:

- **time**: The task time (ms).
- **distance**: The distance from the correct solution.

Aside from training, the amount of data collected was 16 participants $\times$ 4 visualization methods $\times$ 4 levels of element counts $\times$ 4 levels of differences = 1024 trials.

3. RESULTS AND DISCUSSION

The design was analysed as a linear mixed-effects model (within-subjects ANOVA) followed by post-hoc test. The relevant packages in R are *nlme* and *multcomp*, and functions *lme* and *glht*, respectively.

3.1. Time

The grand mean for task completion time was 10.3 seconds (Table 1). The main effects of **type** ($F_{3,45} = 11.2, p < .001$) and **task** ($F_{15,225} = 60.8, p < .001$) were statistically significant. There was a significant **type** $\times$ **task** interaction effect (Fig. 6, Fig. 9), which was due to the baseline condition *table*.

Table 1: Mean of time on task

	Mean of time (s)	%
Overall	10.3	
Stacked bar chart	8.9	0%
Doughnut chart	9.9	+11%
Pie chart	10.5	+18%
Table	11.7	+32%

Figure 6: Interaction plot for **type** and **task**.

Multiple comparisons of means with *Tukey HSD* show that all the other differences except between *pie* and *doughnut* are statistically significant (Table 2).

3.2. Distance

Overall, 957 tasks were performed without errors, 58 were one swap from the correct result, and 9 were

two swaps away. The grand mean for errors was 7.4 (as a rate of swaps occurring per hundred tasks, Table 3). The main effects of **type** ($F_{3,45} = 3.22, p < .05$) and **task** ($F_{15,225} = 60.8, p < .001$) were statistically significant, and there was no interaction.

Table 2: Comparison of task time means

	Estimate Std.	Error	z value	Pr(> z)
doughnut-bar==0	948.5	318.6	2.977	0.0152 *
pie-bar==0	1577.0	318.6	4.949	<0.001 ***
table-bar==0	2543.3	318.6	7.982	<0.001 ***
pie-doughnut==0	628.5	318.6	1.972	0.1985
table-doughnut==0	1594.8	318.6	5.005	<0.001 ***
table-pie==0	966.3	318.6	3.033	0.0133 *

Table 3: Mean of answer distance

	Mean of distance (swaps/100 tasks)	%
Overall	7.4	
Stacked bar chart	8.2	0%
Doughnut chart	9.0	+10%
Pie chart	10.6	+29%
Table	2.0	-76%

The only statistically significant differences (multiple comparisons, *Tukey HSD*) were between *table* and other conditions (Table 4). The counts of false answers (out of 256) are listed in Table 5.

Table 4: Comparison of distance means

	Estimate Std.	Error	z value	Pr(> z)
doughnut-bar==0	0.00048	0.0253	0.019	1.00000
pie-bar==0	0.02099	0.0253	0.828	0.84116
table-bar==0	-0.06640	0.0253	-2.619	0.04388 *
pie-doughnut==0	0.02050	0.0253	0.809	0.85031
table-doughnut==0	-0.06689	0.0253	-2.638	0.04154 *
table-pie==0	-0.08740	0.0253	-3.447	0.00316 **

3.3. Subjective opinions

Table 6 shows how the participant replied to questions 'which visualization technique is the most pleasing/second-pleasing/third-pleasing to use?'

The participants were also asked 'what do you think, which visualization technique is the fastest to understand?' The stacked bar chart got six votes (37%), the pie chart four (25%), and the doughnut chart and table got both three votes (19%).

3.4. Discussion

Fig. 7 shows the means and 95% confidence intervals for task execution times for the four visualization techniques. As noted in subsection 3.1, all the other differences between means of time are statistically significant except between doughnut and pie chart. The results suggest that the doughnut chart might be slightly more efficient than the pie

Table 5: Error count

	Error count	%
Overall	67	6.5%
Stacked bar chart	17	6.6%
Doughnut chart	20	7.8%
Pie chart	25	9.8%
Table	5	1.9%

Table 6: Subjective ranking: pleasantness

	I	II	III	Σ
Pie	10	2	3	15
Doughnut	0	2	7	9
Bar	3	9	3	15
Table	3	3	3	9
Σ	16	16	16	48

chart, but the statistical power of this experiment leaves it inconclusive. One interesting observation is that the serially-processed table is actually *faster* than the pie and doughnut chart with four elements, and the serious slowdown starts as late as at seven elements (Fig. 9).

Figure 7: The mean of task *time* by *type* with 95% confidence intervals.

This performance order of the techniques is in line with the evaluation of elementary perceptual tasks (Cleveland and McGill 1984): the stacked bar chart is based on the estimation of length which is the easiest to perform. It is unclear what the pie and doughnut charts are based on, but it probably is a complex mixture of estimating angle and area. The baseline condition, *table*, is the slowest because of the serial processing. However, the table is the only technique that provides exact numeric values for elements. This shows in the error rate: there were only five false answers out of 256 (Table 5). The overall error rate was low (only 67 errors out of 1024 tasks), which contributed to inconclusive statistical results except between *table* and other conditions (Table 4 and Fig. 8).

The experimental task was chosen keeping in mind the critique faced by the earlier studies. The task involves no estimation of exact numbers (von Huhn 1927), there are no visual computations (Spence and Lewandowsky 1991), and the role of short-term memory is kept at minimum (Peck et al. 2013). The setup is ecologically valid as most guidelines recommend to use this kind of visualization techniques for no higher proportion counts than seven. The appearance of visualizations is as plain as possible, and they were produced with the de-facto standard statistical system R (R Development Core Team 2014).

Figure 8: The mean of *distance* by *type* with 95% confidence intervals.

One possible problem was noticed in the stimuli. We have certain sensitivity to pie chart values that are multiples of 90 degrees, possibly because of the clock face analogy. This makes it more efficient to perceive 25%, 50%, and 75% proportions from a pie chart than other percentages. In light of this, the part of data with 24% difference in stimuli was unfortunate. However, the data shows that the 24% condition did not stand out as an exception.

Subjective opinions about these visualizations provide some insight on why the pie charts are so controversial. Twelve participants (75%) indicated the pie chart either the *most pleasing* (10) or the *second-pleasing* (2) visualization. In addition, almost half of the participants (7 out of 16) perceived either the pie or doughnut chart to be the *fastest* visualization to understand. Only six participants perceived the bar chart to be the fastest technique, and half of them still regarded the pie chart the most pleasing to use.

How do the results of this experiment relate to the claims in the introduction? It seems that the strong dislike of pie charts by distinguished practitioners is not so well-founded as they think. The performance penalty, both in task time (Table 1) and accuracy (Table 3) is there, but is it large enough that we

Figure 9: Interaction plot for **count** and **time**.

should bypass the users' preferences? Based on these results, the selection between bar, pie, and doughnut charts is more a matter of task and style than following some strict doctrine. These results do support the last claim: if we have a small number of items to show, and the actual values have any significance, then there is no reason for not showing the numbers, i.e., a table of them.

4. CONCLUSION

This paper compared three of the most common visualization techniques of proportions in terms of their task execution time, correctness of the outcome, and subjective satisfaction. The result shows that the bar chart is the most efficient technique, followed by the related doughnut and pie charts. The result also suggests that the popular doughnut chart might actually be a better performer than the pie chart, but this remained statistically inconclusive in this experiment, and needs further study. In any case, the cause of the difference might be the distinction in reading (combination of angle and area vs. area).

ACKNOWLEDGMENTS

Thanks to Jorge Ávalos-Salguero for implementing the *GraphTest* application.

REFERENCES

- Barrow, J. (2011). *Cosmic Imagery: Key Images in the History of Science*. Random House.
- Cleveland, W. S. and R. McGill (1984). Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association* 79(387), 531–554.
- Eells, W. C. (1926). The relative merits of circles and bars for representing component parts. *Journal of the American Statistical Association* 21(154), pp. 119–132.

Ehrenberg, A. S. C. (1977). Rudiments of numeracy. *Journal of the Royal Statistical Society* 140(3), 277–297.

Few, S. (2007). Save the pies for dessert. Visual Business Intelligence Newsletter. <http://www.perceptualedge.com/articles/08-21-07.pdf>, visited 2014-06-17.

Few, S. (2013). A pie in the face for information visualization research. Visual Business Intelligence Newsletter. <http://www.perceptualedge.com/blog/?p=1492>, visited 2014-06-17.

Gabrielle, B. (2013). Why Tufte is flat-out wrong about pie charts. Web log post. *Speaking PowerPoint – The new language of business*. <http://speakingppt.com/2013/03/18/why-tufte-is-flat-out-wrong-about-pie-charts/>, visited 2014-06-17.

Kosara, R. (2011). In defense of pie charts. Web log post. *eagereyes*. <http://eagereyes.org/criticism/in-defense-of-pie-charts>, visited 2014-06-17.

Nussbaumer, C. (2011). Death to pie charts. Web log post. *Storytelling with Data*. <http://www.storytellingwithdata.com/2011/07/death-to-pie-charts.html>, visited 2014-06-17.

Peck, E. M. M., B. F. Yuksel, A. Ottley, R. J. Jacob, and R. Chang (2013). Using fNIRS brain sensing to evaluate information visualization interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 473–482. ACM.

Playfair, W. (1801). *The statistical breviary; shewing the resources of every state and kingdom in Europe*. Re-published in Wainer, H. and Spence, I. (eds.), *The Commercial and Political Atlas and Statistical Breviary*, 2005, Cambridge University Press, ISBN 0-521- 85554-3.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <http://www.R-project.org>.

Spence, I. and S. Lewandowsky (1991). Displaying proportions and percentages. *Applied Cognitive Psychology* 5, 61–77.

Tufte, E. R. (1983). *The Visual Display of Quantitative Information*. Graphics Press.

von Huhn, R. (1927). Further studies in the graphic use of circles and bars: A discussion of the Eells experiment. *Journal of the American Statistical Association* 22(157), pp. 31–39.