Technology and Testing

Designing Effective Data Visualizations for Testing

By: Brian D Bontempo, PhD *Mountain Measurement, Inc.*
Brian served on the NCCA Standards Revision Steering Committee and is a Psychometric Reviewer for the NCCA.

The opinions expressed within are those of the author and do not reflect the views of the National Commission for Certifying Agencies or the Institute for Credentialing Excellence.

Introduction

This article aims to synthesize the literature on data visualization design with my own personal experience to provide readers with some guidance when creating data visualizations for testing. Since data visualization design is a large topic, this article serves as a beginning primer aimed at individuals designing visualizations for the testing industry. Future editions may apply these design principles to examinee score reports and testing dashboards.

This is the second article in Technology and Testing dedicated to the topic of data visualization. Readers not yet familiar with data visualization are encouraged to read Bontempo, 2014 to gain a basic understanding.

In April, I had the chance to attend the OpenViz conference in Boston, MA. Many of the conference presenters, which included some respected data visualization designers such as Mike Bostock, John Resig, and Robert Simmon, explicitly divulged their basic data visualization design process. Although these varied, a common theme emerged. Data visualization design is an iterative process in which the product emerges and improves as ideas are explored through a trial and error process. These designers believe that design is as an art rather than a science. Even so, I believe that the basic process for designing explanatory data visualizations can be divided into the following steps,

1. Identify the purpose
2. Choose a chart type
3. Populate with data
4. Add attributes
5. Optimize the layout

Each of these will be explored.

Identify the Purpose

The design of a visualization can be greatly impacted by the purpose, so it is wise to identify the purpose and the intended audience before beginning the design process. Generally speaking, there are three purposes that a data visualization may have, to inform the user, to persuade the user, and to inspire the user to ask more questions of the data.

Often in testing, the purpose of a visualization is to inform the user, such as the visualizations used in examinee score reports. For these, the data is just as important as the visualization. Therefore, the data must be presented from a neutral perspective (Steele and Iliinsky, 2011).

In testing, persuasive visualizations are often used in hand-crafted technical or program management reports. The conclusion is more important than the data for persuasive visualizations, so attributes are often used to illuminate the designer's interpretation of the data.

Both informative and persuasive reports can be provided to users who are interested in taking action on the information provided without having the desire to engage with the data any further. On the other hand, there are many users that are interested in diving deeper into the data (Bontempo, 2012). For these power users, visualizations that inspire and permit the user to interact with the data are required. In licensure testing, educators and regulators are an emerging class of power users whose desire for data may now match their desire for findings. Designers that create interactive data visualizations are likely to succeed with these types of users.

Although interactive design is a very important aspect of data visualization, it is an entirely separate science. Readers interested in learning more about this topic are encouraged to start with Shneiderman and Plaisant's work.

Choose a Chart Type

Once the purpose is clear, the next step is to choose which type of chart to construct. Loosely speaking, there are four types of charts, tables, graphs, maps and a catch all category which I'll call innovative visualizations. We'll explore some reasons to use each type.

Tables are effective at providing numbers in a logical, organized manner. As a result, they are the most useful chart type when the numeric values themselves are important. Therefore, tables are well suited to visualizations that have an informative purpose and are often useful for examinee score reports.

Tables are also a useful way of conveying multiple constructs simultaneously especially when those constructs have different units of measurement (Few, 2004). This makes tables a useful way of conveying information in technical reports where different item performance or test performance metrics are displayed concurrently.

In contrast, graphs are pictures in which the placement of objects in at least one dimension of space spatially conveys quantitative data. Graphs are useful for conveying the relationship between variables or displaying trends. For graphs, the picture and the conclusions one may draw from them are more important than the numeric details. There are many types of graphs, so once a designer has decided to use a graph, the next step is identifying which type of graph to use.

- Scatter plots are a useful way of displaying the relationship between two equal interval or ratio variables. In testing, scatter plots effectively convey the relationship between two subscores of a test.
- Line charts are an effective way of displaying the changes in one or more quantitative variables over time. In testing, line charts are a good way of expressing the test volume or passing rate over time.
- Bar charts are a valuable way of displaying the differences between cross-sectional quantitative values associated with different entities of a categorical variable. In testing, bar charts can successfully communicate the cross-sectional differences in the test volume of different groups.
- Histograms are the gold standard when it comes to visually displaying the distribution of values for one quantitative variable. In testing, the distribution of test scores is validly illustrated as a histogram.
- Box plots are a helpful way of visually summarizing the distributions of multiple quantitative variables. In testing, box plots are a useful way of displaying the cross-sectional performance of different groups of examinees such as those from different jurisdictions or different educational programs.
- Although pie charts are commonly used, they are not a typically an effective way of communicating data since the eye has a difficult time comparing the size of objects when they are placed in a circle. Designers are generally encouraged to use bar charts instead of pie charts. However pie charts can successfully convey valuable information when space is limited. For example, pie charts are effective at conveying a percentage associated with each state on a map such as the passing rate by state.

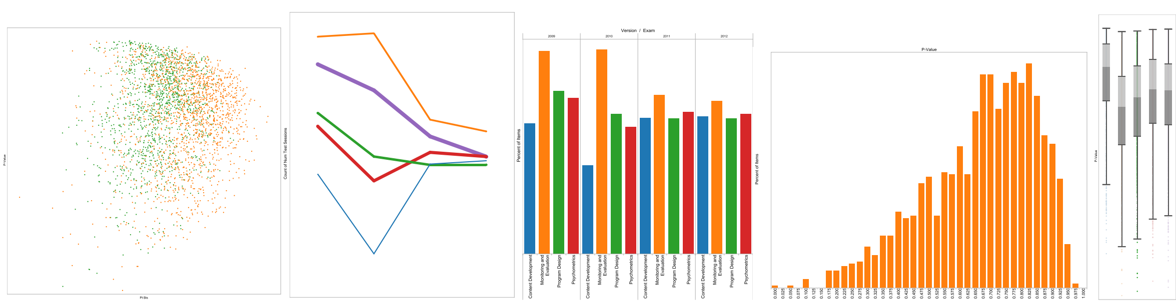

Figure 1. Examples from left to right of a scatter plot, line chart, bar chart, histogram, and box plot.

It should be obvious that maps are used to convey geospatial information. Historically, data visualization maps have been drawn to scale and populated with numeric values or icon sized visualizations. Some, called choropleths (Dupin, C. 1826 and Wright, J. 1938), have added complexity by shading various regions to represent a quantitative variable associated with the region. In recent years, designers have created maps which morph the size of the geographic elements (e.g., states) to match the size of one or more quantitative variables. These maps are called cartograms (Gillard, Quentin, 1979). Although cartograms maps may be fun to create and read, their real utility may be limited for testing.

There are number of innovative visualizations that are being used frequently enough that one may question whether or not these should be considered innovative anymore. These include 3D, heat maps, bubble charts, tree visualizations, word cloud, and many other less common types. Since each has a specific application, designers not satisfied with the more common chart types are encouraged to explore these alternative chart types before creating a new chart type from scratch.

Figure 2. a bubble chart and word cloud displaying the performance by topic for a failing examinee where the large circles and text represent areas where the examinee should focus their study efforts.

This raises an important point when choosing a chart type. Users that are familiar with a chart type, read, understand, and interpret visualizations of that type more quickly and accurately than those unfamiliar with the type. Therefore, designers are encouraged to curb their enthusiasm for innovation and select the common chart types whenever they are appropriate. Keep in mind that most data visualizations in testing are successfully created with traditional chart types.

Populate with Data

Although this step should be self-explanatory, it is worth mentioning that innovative visualizations often require several rounds of selecting a chart type, populating the chart with data, and reflecting upon the creation. It is through this trial and error process that, upon reflection, designers may opt for an alternative data set or a different chart type. Although this is the most common spot for iteration, designers may also find it necessary to repeat these steps after they have begun to add attributes or optimize the layout. Keep in mind that the majority of testing visualizations do not require iteration since traditional chart types succeed in fulfilling their purpose.

At some point, data visualization designers may ask, “How much data should I provide?” The answer to this question varies widely. On one side, in his review of Napoleon’s March to Moscow, Tufte advocates for visualizations that bring together a number of different variables (Tufte, 1983). On the other side, in his advocacy of identity plots and visualizations which are now called Wright Maps (Wilson & Draney, 2000), Benjamin D Wright has suggested that simpler visualizations which present only two variables are more effective. Either way, it is important for a designer to make conscious decisions about the data being provided noting that too much data can muddy the power of the important data while failing to provide enough data can leave the user with unanswered questions.

A visualization is only as good as the data being visualized. Therefore, it is imperative that the accuracy of the data be validated and the efficiency in which that data is processed verified. With testing visualizations, I have also found that it is necessary for the designer to populate visualizations with the minimum and maximum values that are possible. By doing so, the designer can verify that the visualization is rendering properly and providing visible, useful information for the most extreme users.

Add Attributes

From a data visualization perspective, an attribute is any modification to a chart that helps to illuminate the similarities or differences amongst the objects in the visualization. Visualization attributes can be classified into two categories, those that illuminate similarities and differences amongst categorical variables and those that illuminate quantitative variables.

The following attributes are useful with categorical variables, position, color (hue), shape, fill pattern, line style, font, and sort ordering. Designers are encouraged to use these attributes to group all of the data points from one category together or to help the user to group the data points from similar categories together.

With quantitative variables, position, size (length, width, and area), color (intensity), and sort ordering can be used to help the user quickly perceive similarities or differences amongst the data. Designers are encouraged to exercise caution when using color intensity to express the differences in a quantitative variable. Color intensity does not occur linearly and can be impacted by adjacent colors. Designers are encouraged to review Robert Simmon's work on the Subtleties of Color for a more thorough guide on using color effectively. In addition, designers must provide visualizations that are friendly to the color blind by using color blind sensitive palettes or duplicating/replacing color attributes with another attribute.

Trellising, also known as latticing to R users, is a graphing technique that is quite effective at illuminating similarities and differences. When a designer produces a trellis chart, (s)he crates similar charts in a grid like fashion where each chart displays the data associated with a particular category. Each chart maintains the same set of axes which make it easy to visually compare the relationships across charts. The example below shows a set of trellised histograms for the p-values of the items of four different hypothetical tests. By trellising these histograms, it is easy for the user to compare and contrast the distributions of these four assessments.

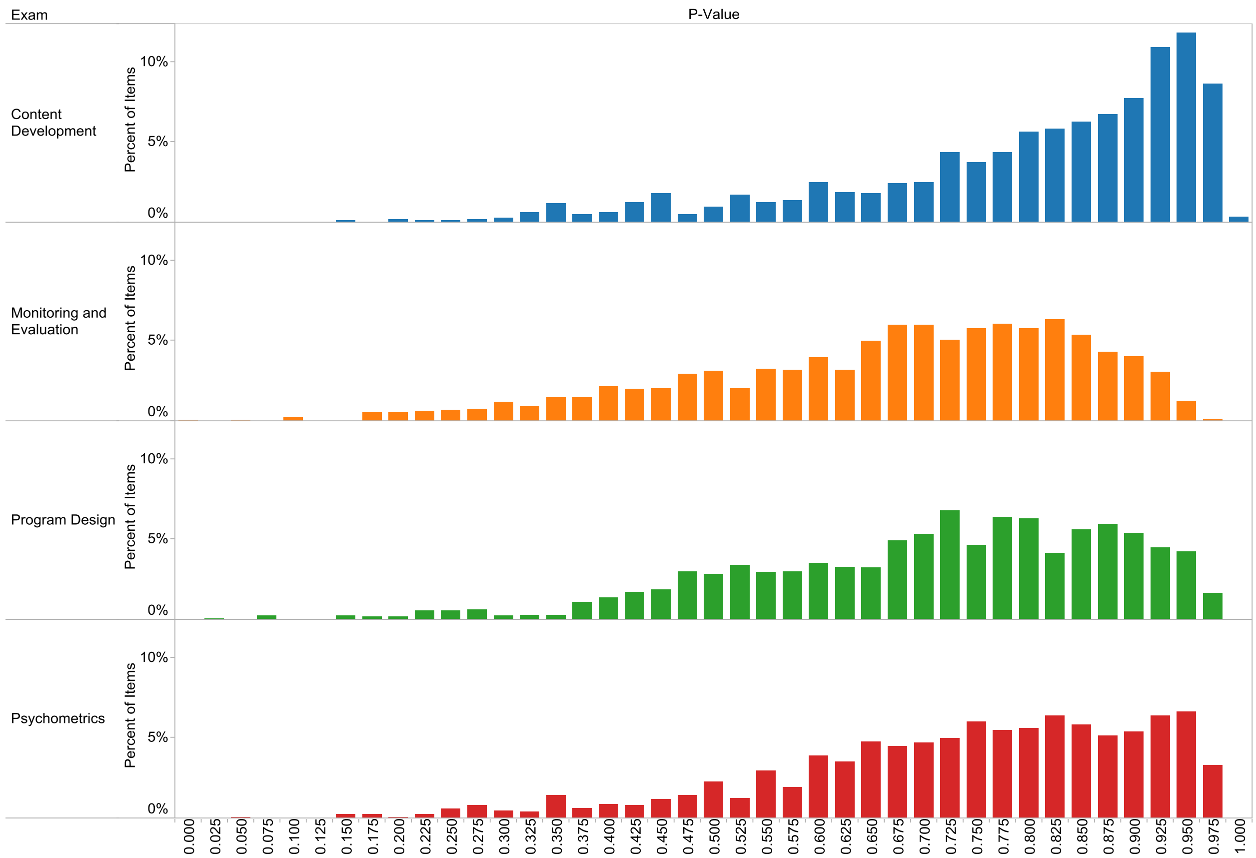

Figure 3. The distributions of the p-values of the items of four hypothetical tests displayed in a trellised manner.

As stated earlier, the purpose of the visualization may largely impact the extent to which attributes should be used. In my experience, persuasive data visualizations in testing do not use attributes enough. Many of the hand-made technical, financial, and program management reports provided by testing professionals could be improved greatly by adding attributes to their visualizations.

Optimize the Layout

When optimizing the layout of a data visualization, it is important to keep this phrase in mind, “Above all else, show the data” (Tufte, 1983). This helps the user efficiently process the visualization without being distracted by the non-essential elements. One strategy that Tufte advocates is maximizing the data ink ratio which can be achieved by minimizing the extent to which chart junk is used. In other words, designers are encouraged to minimize the extent to which gridlines, legends, tick marks, and labels interfere with the data being presented.

Another important consideration in optimizing the layout is to ensure that the visualization has integrity, meaning that the size of the effect in the data should match the size of the effect in the graphic (Tufte, 1983). Although it is tempting to display only the portions of an axis that contain visually perceptible data points, this practice is discouraged since it visually exaggerates any differences in the data. In testing, this is applicable to the reporting of percent scores, where it may be tempting but unwise to exclude values less than 50% from the axis of a visualization.

In addition, traditional graphic design layout principles such as the orientation of labels, text alignment, and number and date formats should be evaluated and adjusted. Designers are encouraged to right align numbers and to maintain the same number of digits following the decimal point. This maximizes the efficiency in which users can perceive these data. Two important graphic design elements that are worthy of attention are column width and row height. These are particularly important to tables but are also applicable to the height or width of the bars found in bar charts. Used effectively, row height and column width can help the user to quickly perceive the direction, horizontal or vertical, in which the data with the most pertinent comparisons should be read.

Concluding Remarks

The five step process of identify the purpose, choose a chart type, populate with data, add attributes, and optimize the layout has worked well for me in creating data visualizations. Sometimes I am provided with a dataset or select the data before choosing a chart type. Since this re-ordering of the steps mimics the iterative process mentioned earlier, designers may wish to consider this ordering as a viable alternative.

Once the visualization is complete, a designer may want to consider adding annotation to further spotlight aspects of the data to the reader. Designers are encouraged to exercise some restraint in their annotations. Subtle annotations say much more than a page filled with non-essential details. If a designer believes that more annotation is required, then I recommend placing explanatory text in proximity to the visualization or embedding the visualization in a written report.

Annotations may be challenging for visualizations that are systematically produced, such as examinee score reports. However, creative algorithms can be created that may provide useful insight into the visualization for novice users.

There is so much more to data visualization design than the basic information provided within this article. Those interested in more information are encouraged to use the references below as a starting point. More importantly, the best way to improve data visualization design skills is through practice.

References

Bontempo, B. (2012). Data Visualization. A presentation at the Annual Meeting of the Association of Test Publishers in Palm Springs, CA.

Bontempo, B. (2014). “Testing in Technology: An Introduction to Data Visualization for Testing.” CLEAR Exam Review, 24 (1): 8-13.

Dupin, C. (1826). Carte figurative de l’instruction populaire de la France.

Gillard, Q. (1979). "Places in the News: The Use of Cartograms in Introductory Geography Courses." Journal of Geography. 78: 114-115.

Simmon, Robert (2014). http://blog.visual.ly/subtleties-of-color-references-and-resources-for-visualization-professionals/

Iliinsky, N. & Steele, J. (2011). Designing Data Visualizations. Sebastopol, CA: O'Reilly Media, Inc.

Shneiderman, B. & Plaisant, C. (2009). Designing the User Interface: Strategies for Effective Human-Computer Interaction (5th ed.). Reading, MA: Addison-Wesley.

Tufte, E. (2001). The Visual Display of Quantitative Information (2nd ed.). Cheshire, CT: Graphics Press.

Wilson, M., & Draney, K. (2000). Standard Mapping: A technique for setting standards and maintaining them over time. Paper in an invited symposium: "Models and analyses for combining and calibrating items of different types over time" at the International Conference on Measurement and Multivariate Analysis, Banff, Canada.

Wright, J. (1938). Problems in Population Mapping, in Notes on Statistical Mapping, with Reference to the Mapping of Population Phenomena. Mimeographed Publication no. 1, ed. J. K. Wright. New York: AGS and Population Association of America, 1–18.