INSTITUTE FOR DEFENSE ANALYSES

Human Factors in Data Visualization, Chapter 7 of NATO SAS-124 Research and Technology Group (RTG), Visualization Design for Communicating Defence Investment Uncertainty and Risk

Paul M. Thompson

January 2023
Approved for public release; distribution is unlimited.
IDA Document NS D-33335
H 22-000502

INSTITUTE FOR DEFENSE ANALYSES
730 East Glebe Road
Alexandria, Virginia 22305

The Institute for Defense Analyses is a nonprofit corporation that operates three Federally Funded Research and Development Centers. Its mission is to answer the most challenging U.S. security and science policy questions with objective analysis, leveraging extraordinary scientific, technical, and analytic expertise.

About This Publication

This work was conducted by the IDA Systems and Analyses Center under the Central Research Program (C7229), "NATO Panel Report." The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgments

Thank you to John W. Bailey, Nancy M. Huff, and Heather M. Wojton for performing a technical review of this document.

For More Information

Paul M. Thompson, Project Leader
pthompso@ida.org, (703) 575-4661

David E. Hunter, Director, Cost Analysis and Research Division
dhunter@ida.org, (703) 575-4686

Copyright Notice

Rigorous Analysis | Trusted Expertise | Service to the Nation

Chapter 7 – HUMAN FACTORS IN DATA VISUALIZATION

Paul M. Thompson, PhD
Institute for Defense Analyses
United States of America

7.1 INTRODUCTION

The human cognitive system is capable of quickly absorbing and processing large amounts of information when it is presented visually. This visual processing ability contrasts with the slower, serial processing that is necessary to understand information contained in tabular datasets [8][9][16][17]. Because of this difference, visual representations of data can be extremely useful in communicating and analysing the information content of datasets. However, the design of data visualizations is critical in conveying information content. The effectiveness of a data display depends on how well it leverages the key human factors that affect one's ability to visually analyse data.

This chapter describes some of the basic features and limitations of human visual cognitive systems that affect the processing of visual data. It offers some insight into the question of how to leverage these characteristics to design data displays that support two important scenarios in defence planning and management:

1. Inform or communicate information (e.g. present results to decision makers)
2. Analyse and identify patterns and parameter relationships in datasets (e.g. perform exploratory data analysis)

In the first scenario, effective data displays must quickly and effectively convey information within the available attention spans of decision makers, who may be receiving enormous amounts of information from numerous sources each day. In the second scenario, the analyst must be able to efficiently and accurately identify key data patterns, parameters, and their interrelationships in large datasets. This chapter provides important insights into the human factors aspects of data visualizations that support each of these scenarios.

7.2 IMPORTANCE OF HUMAN FACTORS IN DATA PRESENTATION

From a cognitive psychology perspective, why would the well-known dictum, "A picture is worth a thousand words," be true? The answer is that visualization creates a message that is absorbed visually almost instantaneously, before the conscious mind can begin to process the information content. For example, consider the graphic in Figure 7-1. Before the reader has read and processed the two words in the graphic, a visual message has come across. This message could perhaps be expressed as "Somebody didn't think this one through very well!" The humour in the graphic stems from the cognitive dissonance between the visual message and the information content of the text itself.

Figure 7-1: A Graphic with Cognitive Dissonance

The distinction between the visual message and its information content is a critical conceptual element in good visualization design. In effective visualizations, the visual message is congruent with–and therefore highlights–the information content of the underlying data.

For example, consider the dataset shown in Figure 7-2, extracted from [1]. This dataset, Anscombe's Quartet, has four statistically equivalent datasets. The same regression line describes each dataset equally well, so that from a statistical standpoint, the datasets are indistinguishable.

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

Four sets of paired (x,y) data. For EACH set:

Mean, variance of the x values = 9.00, 10.00

Mean, variance of the y values = 7.50, 3.75

Equation of the least-squares regression line is:
y = 3.00 + 0.500x

Sums of squared errors (about the mean) = 110.0

Regression sums of squared errors
(variance accounted for by x) = 27.5

Residual sums of squared errors
(about the regression line) = 13.75

Correlation coefficient = 0.82

Coefficient of determination = 0.67

Figure 7-2: Anscombe's Quartet

An analyst searching for patterns in the data or looking to communicate such patterns to decision makers might be tempted to stop here and conclude that the datasets come from identical populations. However, as Figure 7-3 shows, a visualization of these datasets shows that this conclusion would be an error.

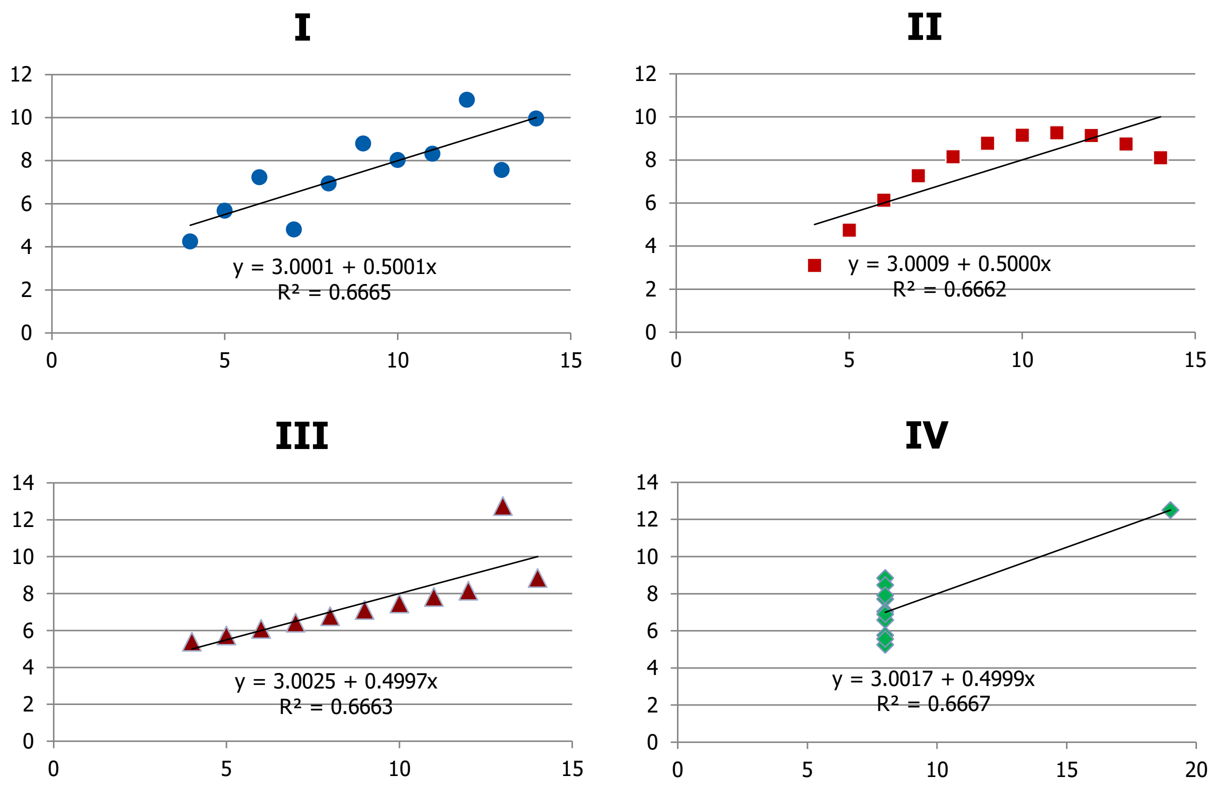

Figure 7-3: Graphs of Anscombe's Quartet

Visually, it is clear at once that the four datasets come from very different populations. This insight is key: if presented properly, human visual systems can acquire and process a large quantity of information at a glance. Skilled analysts use this feature of human visual processing to enhance the message they wish to convey.

7.3 PREATTENTIVE VISUAL ATTRIBUTES

The visual attributes that the human visual system detects and processes very rapidly and accurately, before the cognitive system even has time to focus on them, are called *preattentive* attributes. Perception of these attributes is also largely involuntary: you can't *not* notice them. On a large multi-element display, preattentive tasks are generally performed in less than a fifth of a second, whereas for non-preattentive attributes, much slower attentive processing (e.g., serial search) is necessary. However, a combination of preattentive attributes is not preattentive! [17].

Preattentive visual attributes fall naturally into five general categories: colour, form, spatial aspects, frequency, and motion. Figure 7-4 lists preattentive attributes for each category and gives some indication of the data types for which they are most useful in data displays.

Colour	Shape	Spatial Aspects	Frequency	Motion

Hue **Intensity** Lustre Contrast	**Orientation** **Closure** **Curvature** **Intersection** **Terminators**	**Length, Height, Width** **2-D position** **Proximity** Angle Area/Size Lighting direction Apparent depth	**Density** **Number (estimated)**	**Speed** Direction
	Key: **Green**: most useful for categorical data **Blue**: most useful for categorical data, and sometimes for ordinal data **Purple**: most useful for numeric data (interval or ratio data) **Black**: not as useful as other attributes			

Figure 7-4: Some Preattentive Attributes

When considering how to best design a visual display of data, it is helpful to leverage the ability of human visual systems to quickly process the above-listed preattentive attributes. In general, the effectiveness of attribute choice depends on data type. Some preattentive attributes are especially good for categorical data–for distinguishing and grouping objects. Others are especially good for interval or ratio data–for estimating or comparing numerical values. This section focuses on a few of the more useful preattentive attributes for data visualization. It provides examples of how they may be used effectively or ineffectively, and discusses how the resulting visual displays either help or hinder receipt of the desired message.

7.3.1 Colour and Shape

The human visual system preattentively groups things together by colour and by shape. Consequently, these two preattentive attributes are useful in conveying categorical data. Because these attributes are effective only if the set of colours or shapes used are readily distinguishable, they are most useful in distinguishing a few, rather than many, categories. For example, in Figure 7-5, it is immediately apparent that there are two colours and two letter shapes.

Figure 7-5: Hue and Shape Convey Categorical Data

The perception of colour is not uniform across the human population. For instance, about 8% of males and less than 1% of females have red-green colour blindness, the most common form [12]. Therefore, including both red and green in a display's colour set should be avoided, especially when these colours appear in small markings, which increases the difficulty for some viewers to distinguish them.

Colour blindness concerns can be mitigated by combining hue with another preattentive attribute such as shape or colour intensity. For instance, in the radar sensor plot in the left panel of Figure 7-6, a viewer with normal vision easily distinguishes two colours (red and green), along with three shapes (vertical bars, dashes, and pluses). This preattentive processing detects two colour categories (corresponding to green and red), and three shape categories (corresponding to bars, dashes and pluses). However, a viewer with red-green colour blindness would preattentively detect only the three shape categories. In both cases, the viewer preattentively processes the categorical information on the sensor plot, but the effect is stronger for people with normal vision.

The radar track picture in the right panel uses two additional colours to distinguish Identification, Friend or Foe (IFF) Mode 2 (yellow) and unique Mode 4 (cyan). In both Figure 7-5 and Figure 7-6, notice that the preattentive effect of the colour difference is stronger than the preattentive effect of the shape difference. Viewers with normal vision or with red-green colour blindness will preattentively register the presence of these two IFF modes on the display.

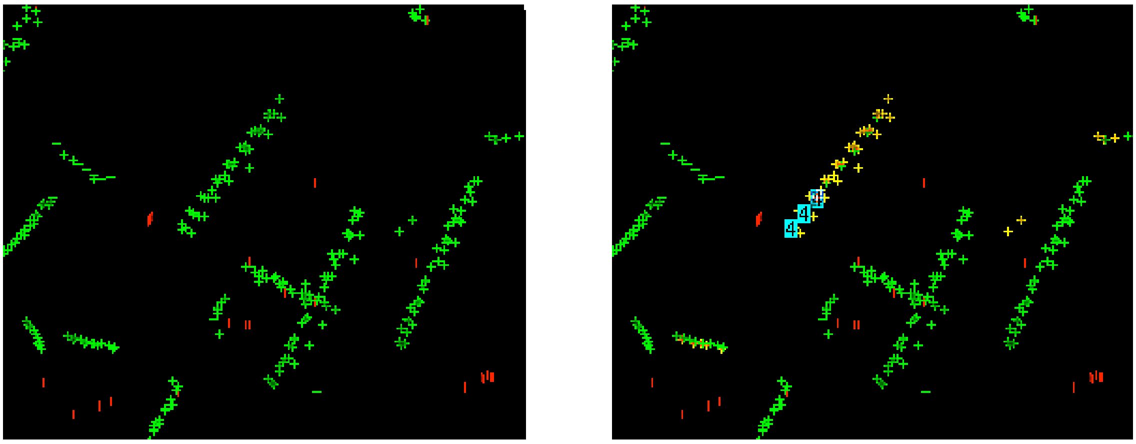

Figure 7-6: Grouping by Similarity: Communicating Friendly IFF/SIF Information
Image: Summers, Scott H., 2010, "Preattentive Attributes in Visualization Design: Enhancing Combat Identification," Chapter 8 in Andrews, Dee H., Robert P. Herz, Mark B. Wolf, *Human Factors Issues in Combat Identification*, CRC Press. © Dee H. Andrews, Robert P. Herz and Mark B. Wolf 2010. Reproduced with permission of the Licensor through PLSclear.

Colour can be a good choice for indicating ordinal data (rank of object groups) if the colour scheme follows a natural or easily understood gradient. For instance, Figure 7-7 uses four levels of increasing intensity of the same hue (blue). The visual system can easily (and preattentively) understand colour intensity as encoding ordinal data.

Figure 7-7: Colour Intensity is Useful with Ordinal Data

The rainbow spectrum of colours can also effectively encode ordinal data. This ordered set of colours that people commonly know from elementary school years is not useful for red-green colour-blind people.

The risk matrix shown in Figure 7-8 illustrates use of the common "stoplight chart" RYG (red-yellow-green) colour combination to group outcomes into three levels of increasing risk. The visual cognitive system preattentively groups the matrix elements into three ordered groups by colour, and only then begins to process the other elements in the graphic.

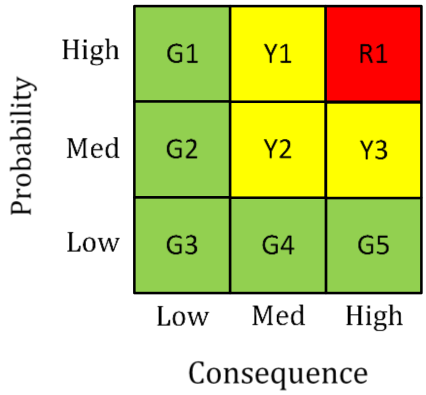

Figure 7-8: Hue Can Be a Good Choice for Encoding Ordinal Data

If not used carefully, however, hue can be ineffective in encoding ordinal data. For example, if yellow indicates the grouping with lowest risk, and green the grouping with highest risk in the risk matrix, then the rank ordering would be yellow < red < green. This usage of colour would create cognitive dissonance with the commonly used, normal stoplight chart schema, and as such would not be a good choice for preattentive processing for ordinal data.

For similar reasons, the set of colours that the evapotranspiration chart in Figure 7-9 uses does not preattentively indicate the ordinal relationships contained in the underlying data. The sequence of colours doesn't entirely follow the natural rainbow order, so attentive processing is necessary. Moreover, the abrupt and arbitrary change in luminance between 0.79 and 0.80 creates an apparent East-West regional split. The preattentively acquired visual effect is that there are two main groupings (East and West, based on luminance) and that within each of these, there are several subgroups (based on hue). The preattentively perceived luminance-based split is artificial; it does not match the evenly spaced gradient of numeric values on the evapotranspiration scale. Although the set of colours satisfactorily indicates the categorical nature of the data, attentive analysis is necessary to satisfactorily process the ordinal nature of the displayed dataset.

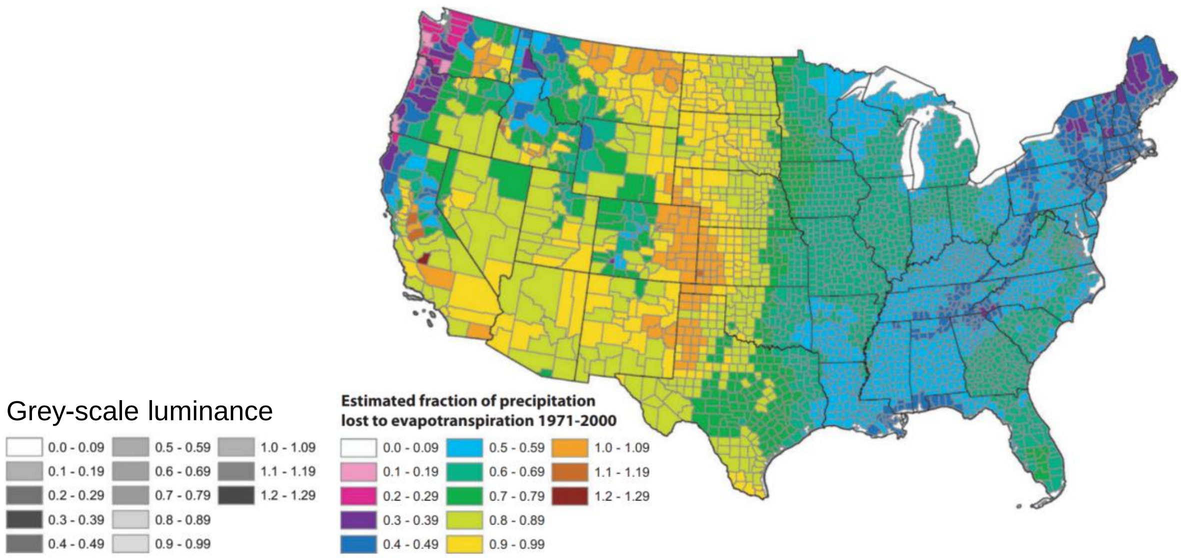

Figure 7-9: Hue Can Be a Poor Choice for Encoding Ordinal Data

Map source: Ward E. Sanford and David L. Selnick, "Estimation of Evapotranspiration across the Conterminous United States Using a Regression with Climate and Land-Cover Data" (2013). Luminance data source: https://eagereyes.org/basics/rainbow-colour-map, accessed 30 June 2014.

Colour is generally not effective for denoting numeric – interval or ratio – data, as Figure 7-10 illustrates. It is impossible to preattentively determine the numeric value–whether on an interval or ratio scale–that corresponds to each data block. Attentive, serial processing is necessary, but this serves only to estimate the underlying numerical values.

Figure 7-10: Colour is Ineffective for Conveying Numeric Data

7.3.2 Proximity

Proximity is a spatial dimension that is useful for preattentively grouping items by category. Figure 7-11 illustrates this characteristic: the eye preattentively sorts the figure into three rows and three columns. Only later does the viewer notice that the rows and columns are made up of five squares, and that there are five different colours of squares. For the human visual system, things that are closer together get grouped together!

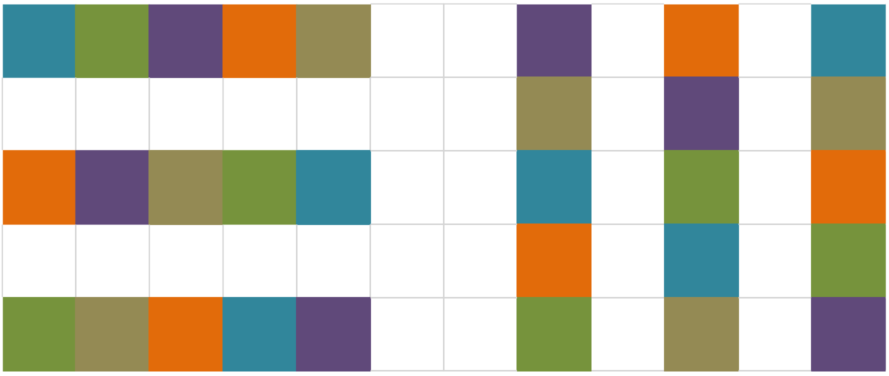

Figure 7-11: Preattentive Grouping by Spatial Proximity

For graphs of interval or ratio data, proximity can be misleading. Since the preattentive visual system captures the minimum distance ("nearness") between curves, not the vertical distance between them, it does not correctly interpret superposed curves. For example, in Figure 7-12, the vertical difference between the two curves has a constant value of 8.0. However, because the nearest distance between the curves decreases from left to right, the visual system erroneously interprets this to mean that the corresponding functions become closer in value. Correctly estimating the vertical distances between these curves requires difficult, slow, attentive, serial processing.

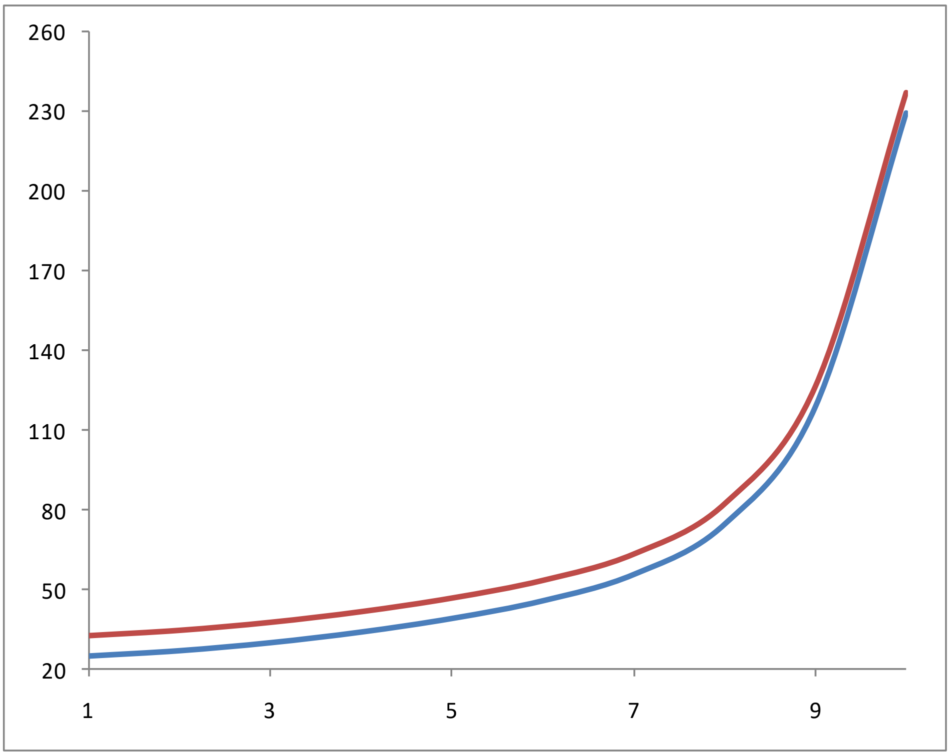

Figure 7-12: Superposed Curves

7.3.3 Position and Distance

On a flat visual display, such as a computer screen or a sheet of paper, two-dimensional (2-D) position is a quantitatively perceived, preattentive visual attribute. Just by glancing briefly at a 2-D visual graphic, a viewer can immediately tell the relative position of the various items that appear in the display. For example, if a viewer wants to identify the positions of the orange squares in a graphic such as Figure 7-11, a brief glance is sufficient to identify where they are located.

Similarly, the human visual system preattentively and accurately compares the distance between pairs of graphic items in a display, such as the individual data markers, reference elements, and x and y axes in Figure 7-3. As a result, 2-D position and distance are very effective in conveying numeric data, whether it be on an interval or ratio scale. Scatterplots and line graphs are commonly used graphic structures that leverage these two attributes.

7.3.4 Length, Height, and Width

Length, height, and width are preattentive visual attributes that precisely convey quantitative data. For this reason, bar charts and column charts are widely used to display numeric information. Indeed, if the purpose of a display is to convey numerical values or differences for interval data or ratio data, bar charts and column charts are often by far the most effective choice, even if they may seem boring from overuse.

7.3.5 Angle Perception

Geometric angle is a preattentive attribute that is quantitatively perceived, but not as accurately as are other visual attributes such as length, height, width, or distance. Because of its relative inaccuracy, it is best to use this “not-so-precise” attribute to display quantitative information only when more “precise” attributes are not available. For instance, comparing angles in a pie chart is not as exact as comparing the height of columns in a column chart; column charts and bar charts are therefore better than pie charts for displaying quantitative information. Figure 6-15 and Figure 6-18 illustrate this point. It is a preattentive task to compare the lengths of the bars in Figure 6-15. The eye immediately notes the monotonic gradient of bar heights and their relative sizes. However, when comparing the pie chart angles of Figure 6-18, there really is no preattentive visual message, except that there are a lot of wedges of different sizes. It takes careful, serial analysis to begin to make sense of the mishmash of colours and wedge sizes.

In the human visual system, patterns of change are most easily observed and compared when the slope of a line in a graph is roughly 45° [3]. Consequently, graphics that change scale so as to “bank” the slopes of lines to 45° maximize the preattentive perception of change. The area charts in Figure 7-13 and Figure 7-14 illustrate this effect.

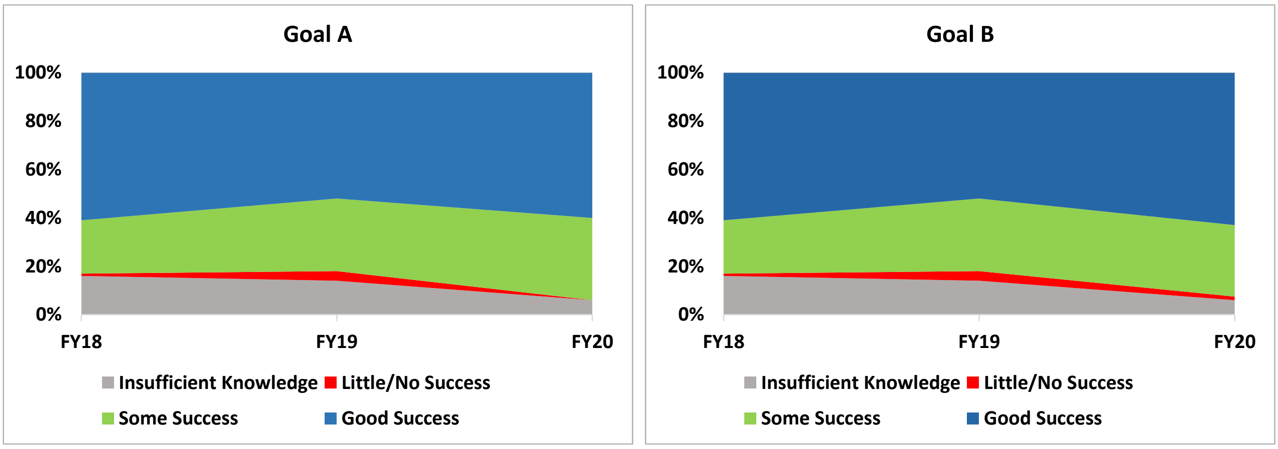

Figure 7-13: Perception of Change Depends on Angle Comparison

In Figure 7-12, the lines appear roughly flat, and it is not easy to tell by how much the slopes of the line segments between the blue and green areas differ in the two panels. Figure 7-13 presents the same information but with the horizontal scale reduced so as to “bank” the slopes of these line segments between the blue and green areas to roughly 45°. Here it is much easier to see by how much they differ.

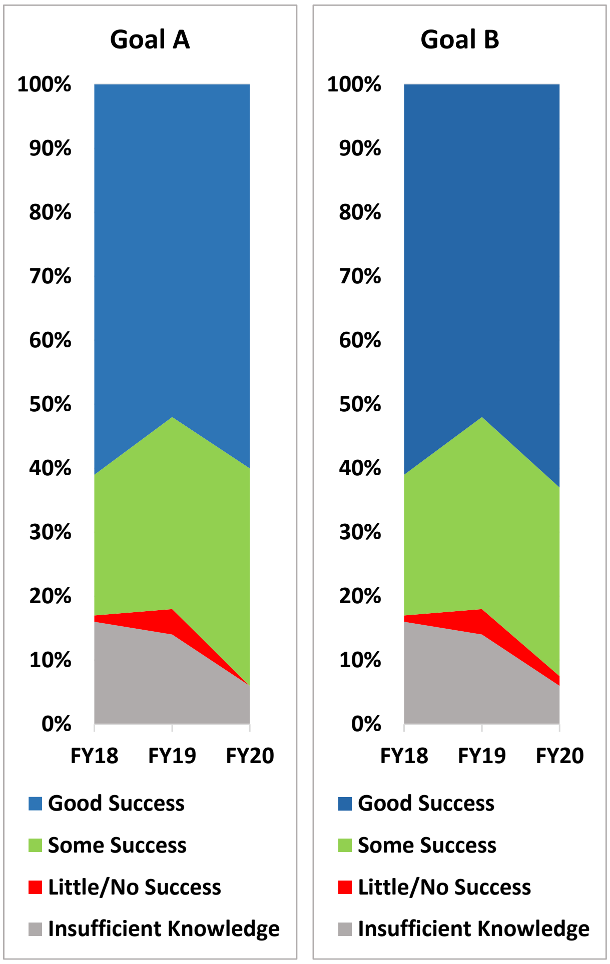

Figure 7-14: Banking to 45° Maximizes Perception of Change

7.3.6 Area Perception is Not Preattentive

For the human visual system, perception of area or size is not preattentive. While one can preattentively see which of two areas is larger if the difference is relatively large, it is not possible to do so if the difference is relatively small. Moreover, it is difficult to estimate by how much two areas differ. For this reason, area should be used to display quantitative information only when a more "precise" attribute (such as height or length) is not an option.

Figure 7-15 illustrates this principle. It is not preattentively obvious how the areas of the four coloured rectangles compare. Moreover, even if you know that the orange rectangle is larger than the other three, which all have the same area, it is not preattentive to recognize how much larger it is (4% in this figure). Because of this difficulty comparing areas, interpreting graphics that use area to encode quantitative data (e.g., multi-pie charts, sand charts) can be onerous.

Figure 7-15: Area Perception is Not Preattentive

Figure 7-8 offers another example of this point. Whereas one preattentively perceives that the different states are divided into administrative areas ("counties" or "parishes") with vastly different areas, it not easy to estimate what the relative size difference is between two of these areas, even in the same state.

7.3.7 Density and Number

Density and number are data attributes that are quantitatively perceived by the human visual system but have limited usefulness.

In the jittered scatterplot in Figure 6-26, for instance, it is easy to see where the density is greatest but hard to say how much denser it is at a given place in the graph compared to another. In Figure 7-8, it is easy to see where the counties are more densely packed, but it is hard to tell what the density difference is.

In graphics that show a small number (up to four) of categories, or a small number of objects within a category, the number is preattentive [13]. This process of instantly counting small numbers is called subitizing [10]. For example, the human visual system quickly notes that there are three bars and three columns in Figure 7-11, two lines in Figure 7-12, two graphs in Figure 7-13 and in Figure 7-14, and four coloured boxes in Figure 7-15. But it does not preattentively recognise that there are thirteen colours in Figure 7-9, no that there are six squares of each colour in Figure 7-11.

7.4 HUMAN VISUAL PROCESSING: ATTENTION

Attention is another key attribute of the human visual processing system. In the context of data visualization, attention is the ability to visually "attend" to an element of the visual field, i.e., to place the mind's focus of attention on that element. Limitations on human visual attention play a significant role in the effectiveness of visual data displays.

7.4.1 Change Blindness

A characteristic of the human visual system is that, when presented with a visual field that has multiple items, it can notice only a single change at a time. The critical factor is attention: to notice a change in an element in a visual field, it is necessary to attend to it. This characteristic results in what is known as change blindness [5].

Change blindness occurs, for instance, as the eye moves and focus shifts from one part of a visual field to another. If a change occurs to an element within the field during the eye movement, one will notice it only if one's attention is focused on that element. Other changes will not be noticed.

Change blindness also occurs if there is a momentary break in the visual field. For example, a bright flash of light can cause change blindness, as will a blank slide in a slide show or slide presentation.

Figure 7-16 illustrates how change blindness can make it difficult to interpret a graphic. The figure shows the results of a before-and-after opinion survey of twelve participants. In order to identify which of the participants' "Before" opinion differs from their "After" opinion, it is necessary to move the eye's focus from the left half to the right half of the graphic. This motion swamps the visual field, thereby causing a failure of the system that allows preattentive perception of change. Hence, the viewer has to use much slower attentive processing that requires moving the eyes' focus back and forth for each of the participants, to observe that exactly six of them changed their opinion.

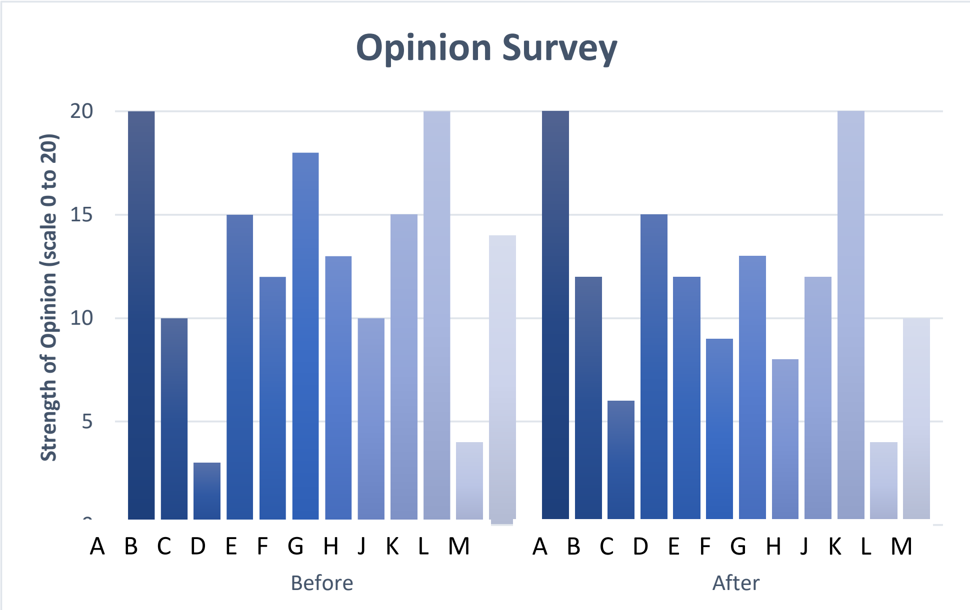

Figure 7-16. Change Blindness

Figure 7-17 presents a different view of the same data. Aligning the "Before" and "After" columns next to each other for each participant eliminates the need to move the eye from one side of the graphic to the other to compare bar heights. This format mitigates the change blindness of Figure 7-16. The preattentive visual system can then quickly identify opinion changes by checking for differences in the heights of adjacent columns.

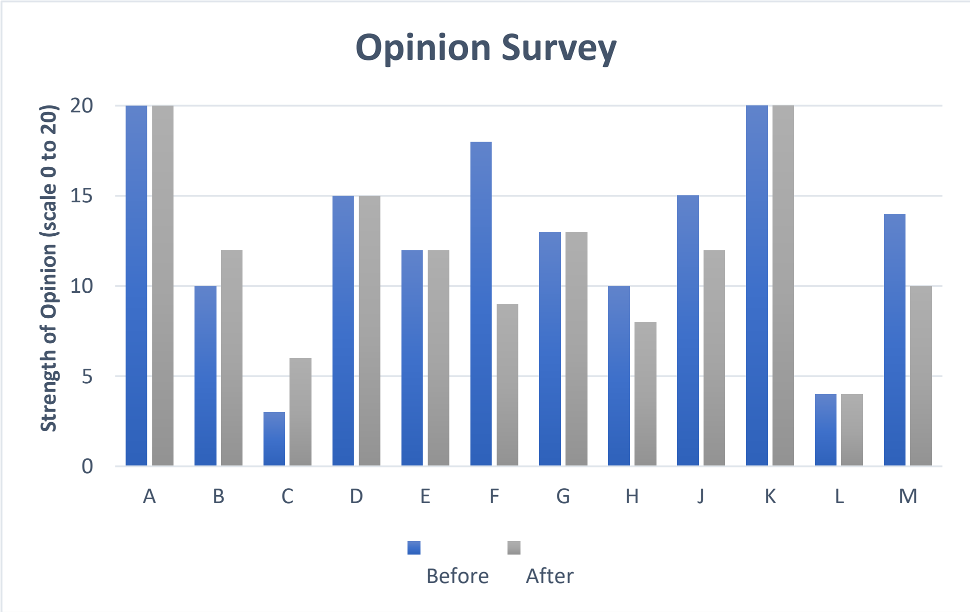

Figure 7-17: Mitigating Change Blindness

7.4.2 Working Memory

Working memory is a cognitive system within the human brain that allows a small amount of information to be held in mind and manipulated during cognitive activity. Its capacity of three to five elements at a time imposes severe limitations on processing ability [3].

Figure 7-18 illustrates why it is important to keep this limitation in mind when designing visual displays. This bar chart shows the results of a two-option (i.e., Low and High) satisfaction survey of participants in twelve U.S. states. A viewer wishing to compare poll results from two states has to keep in mind the states' colours, poll options, poll option locations, and poll results. For example, to visually compare the number from Georgia who polled "Low" with the number from Nebraska who polled "High," a viewer needs to consider twelve elements of information:

- Georgia – darker blue – low – upper part of graphic – second bar from the top – 15
- Nebraska – light orange – high – lower part of graphic – second bar from the bottom – 10

The total number of elements is six per state, or twelve in total for the two states—well beyond the limits of working memory. For this reason, this graphic is not effective in supporting this particular kind of analysis.

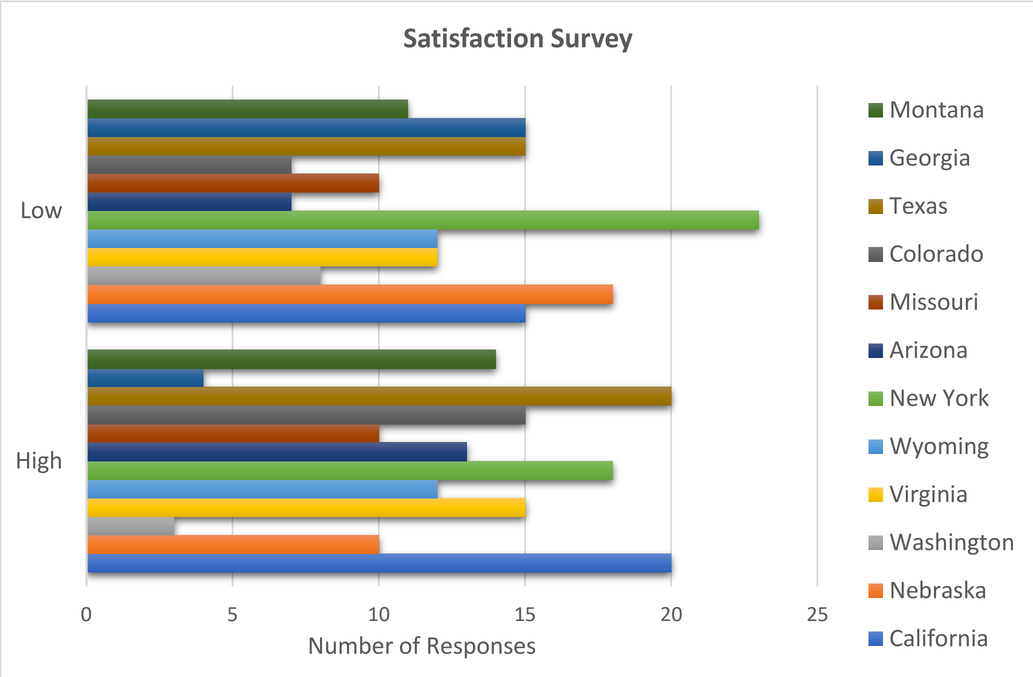

Figure 7-18: Working Memory Limits Visual Processing

In the bar chart in Figure 7-19, several design elements reduce the demand on working memory, for the same comparison of poll results. First, because state names appear directly to the left of the corresponding rows, the viewer does not need to keep them in working memory, only the respective bar locations. Second, unlike in Figure 7-18, the states are ordered alphabetically, making it easier for the viewer to identify the correct rows for comparison. Third, for each state, the "Low" bar appears below the "High" bar. Because low-to-high is the natural order, the viewer doesn't need to keep the bar colour meanings in working memory.

With these design changes, the viewer needs to consider only three elements of information per state:

- Georgia rows location – lower bar – 15
- Nebraska rows location – upper bar – 10

This is a much easier cognitive task.

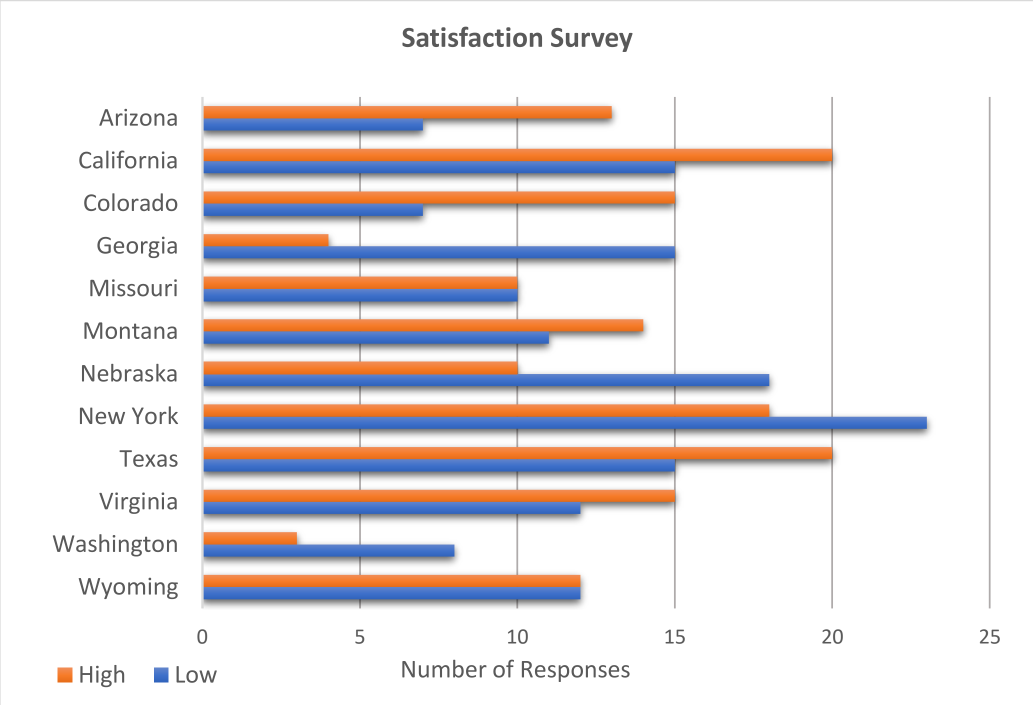

Figure 7-19: Graphic Design Can Mitigate Working Memory Limitations

Although Figure 7-19 is more effective than Figure 7-18 at supporting the particular analytical question posed here, it is not better suited for other questions. For example, suppose the task is to estimate the overall average and variation within each poll result category ("High" or "Low"). This task is easier to do with Figure 7-18, than with Figure 7-19. In general, the effectiveness of a graphic design depends on the analytical question or questions of interest. For this reason, software packages that allow quick and easy customising of graphical displays are highly useful in exploratory data analyses, where analytical questions may not be well defined ahead of time.

7.4.3 Visual Distraction

Visual distraction is a term that is used to indicate the presence of extraneous or redundant elements in a data display. Distraction can inhibit preattentive processing by unnecessarily tying up portions of working memory. Perhaps its strongest effect, however, comes from viewer's not knowing, *a priori*, which—if any—of the graphic elements are not needed to understand the graphic. Viewers must sort through all the elements, determine if distractions are present and identify and mentally discard them. Only then can viewers begin the real process of understanding the data in the graphic. This task is not preattentive

Figure 7-20 illustrates visual distraction. It shows the percent of total college enrollment in the United States that is composed of people aged 25 and older, in the time period 1972 through 1976. To display five data points, the graphic uses four colours; a three-dimensional effect with converging, diagonal lines; a broken vertical axis; and a set of four complex, smooth curves.

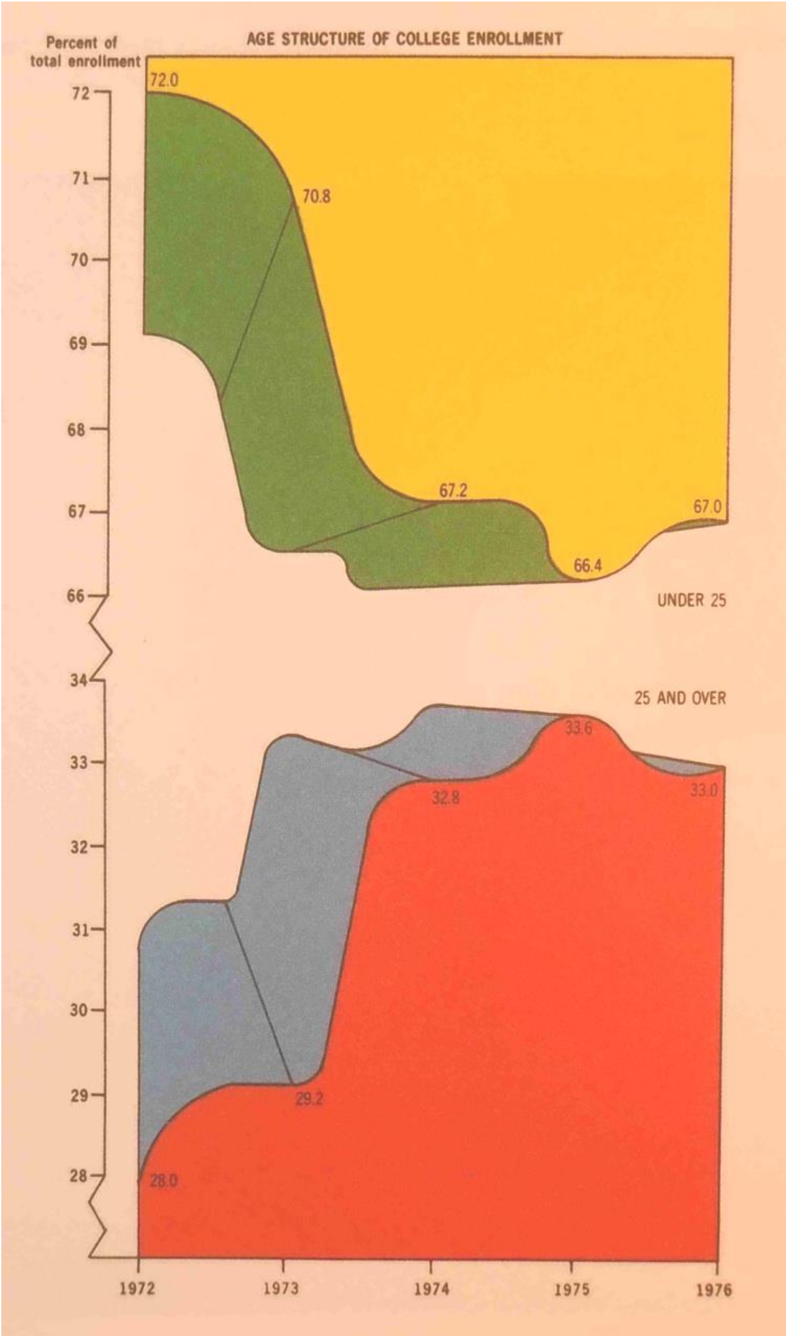

Figure 7-20: Percent of Total College Enrollment Aged 25 and Older as a Function of Time

Graphic source: Wiggins, Rhonda, 1977, "Statistic of the Month: Age Structure of College Enrollment," American Education 13:7, 34, Aug/Sep.

Because of the visual distraction, understanding the data that underlie the graphic is not a preattentive activity. Instead, it takes considerable mental effort to understand the data. The bar chart in Figure 7-21 shows the same data without the distraction, in a manner that allows for rapid, preattentive processing.

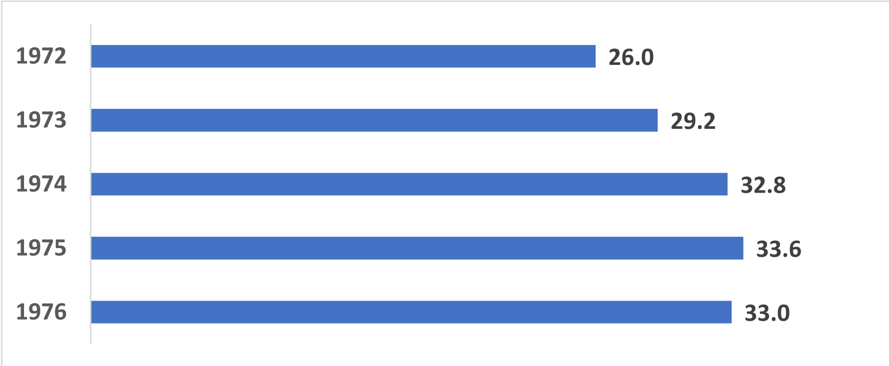

Figure 7-21: Percent of Total College Enrolment Aged 25 and Older

7.5 HUMAN VISUAL PROCESSING: PHYSIOLOGICAL LIMITATIONS

Although the human visual system is able to quickly absorb information, it is also limited in several important ways due to its physiology and structure and the physical nature of light. This section discusses several of these limitations that apply to data visualization, illustrating how they affect the choice and design of effective data visualizations.

7.5.1 Lateral Inhibition

Lateral inhibition refers to the human visual system's innate tendency to emphasizes transitions, processing differences in values rather than absolute values at the boundaries of areas [19]. For example, if a dark area is next to a light area, the dark area next to the boundary will appear darker than it really is, and the light area next to the boundary will appear lighter than it really is. In other words, bright surroundings make an area appear darker than it actually is, and dark surroundings make an area appear lighter than it is. This matter needs to be considered when designing data visualizations.

The column chart in Figure 7-22 illustrates the effect of lateral inhibition. It shows a set of Mach bands–a series of adjacent bars of progressively increasing darkness. Notice that the left side of each column appears to be darker than the right side, even though each bar is coloured uniformly. This effect is due to lateral inhibition.

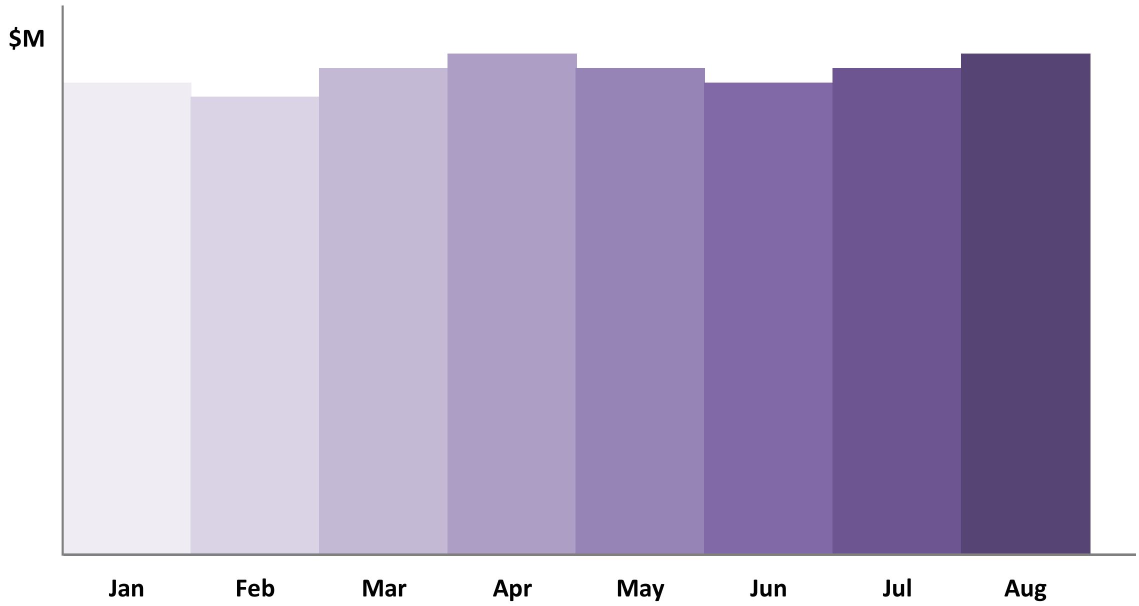

Figure 7-22: Mach Bands Exemplify Lateral Inhibition

Lateral inhibition also influences colour perception. For example, in Figure 7-23, the February and September bars are exactly the same colour. However, since the February bar has dark bars on each side, it appears lighter than it actually is. Similarly, the September bar has relatively light bars on each side, so it appears darker than it actually is. The result is that the viewer perceives the two bars as different in colour, even though their colours are identical.

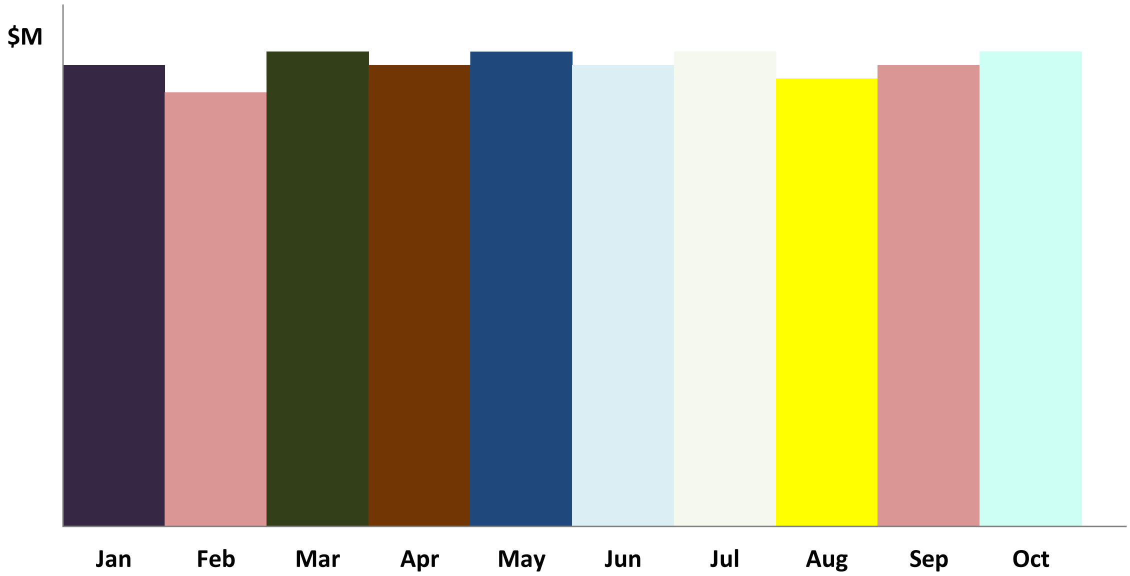

Figure 7-23: Lateral Inhibition Affects Colour Perception

This effect is also apparent in Figure 7-9, where the difference in overall luminance between the eastern half and the western half of the figure makes it difficult to compare the colours of western green counties with eastern green counties.

Lateral inhibition can also cause visual distraction, as the "Hermann Grid" in Figure 7-24 demonstrates. Here, the junctions of the white vertical and horizontal lines are surrounded by white in all four directions, whereas the white line segments between junctions have dark colours on two sides. Because of this, lateral inhibition causes the visual system to perceive the junctions as being darker than the segments. Hence, they appear grey, rather than white. Due to the internal structure of the human eye, this effect is more pronounced on the periphery than at the centre of the field of view.

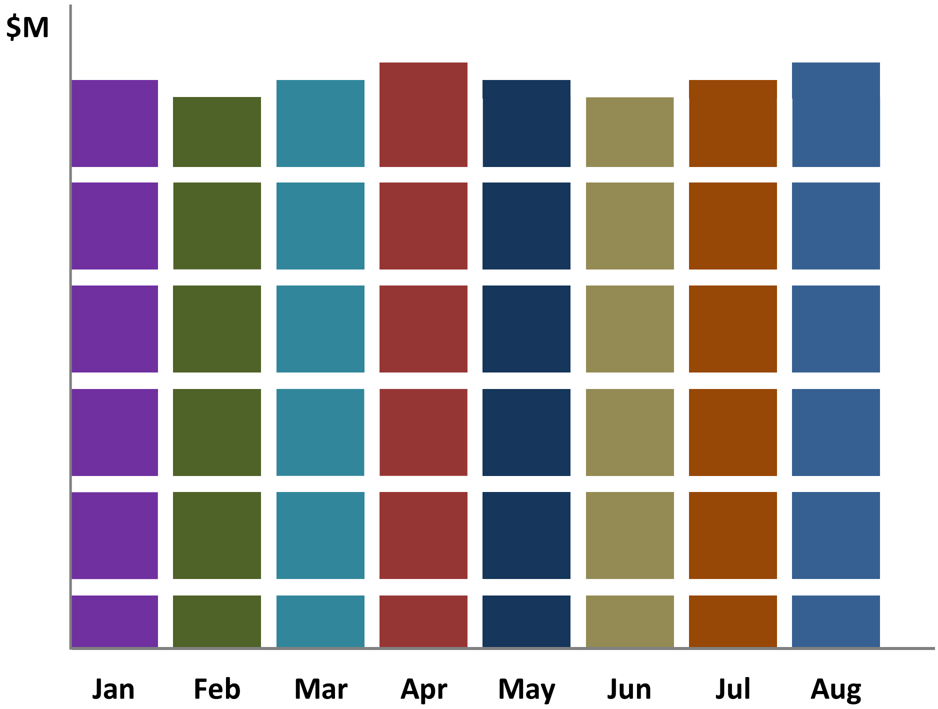

Figure 7-24: Hermann Grid: Lateral Inhibition Can Cause Visual Distraction

In designing a visualization for a given set of data, one can use lateral inhibition to produce a pleasing effect that does not distract from the visual message, such as the graphic in Figure 7-22. In other cases, however, lateral inhibition can cause visual distraction or be misleading, as Figure 7-23 and Figure 7-24 illustrate. Two strategies can avoid or lessen these effects:

(1) Taking care in the choice of colour and luminance for adjoining visual elements.

(2) Providing sufficient neutral space between visual elements so that the effect is no longer noticeable.

7.5.2 Edge Effects

Another limiting feature of the human visual system is its imperfect processing of the edges of elements in the visual field [11]. Although the effect is not noticeable most of the time, the result can be visually distracting. For example, in Figure 7-25, even though each of the columns in the figure is a perfect, vertical rectangle, they appear quite distorted. This apparent distortion is due to edge effects.

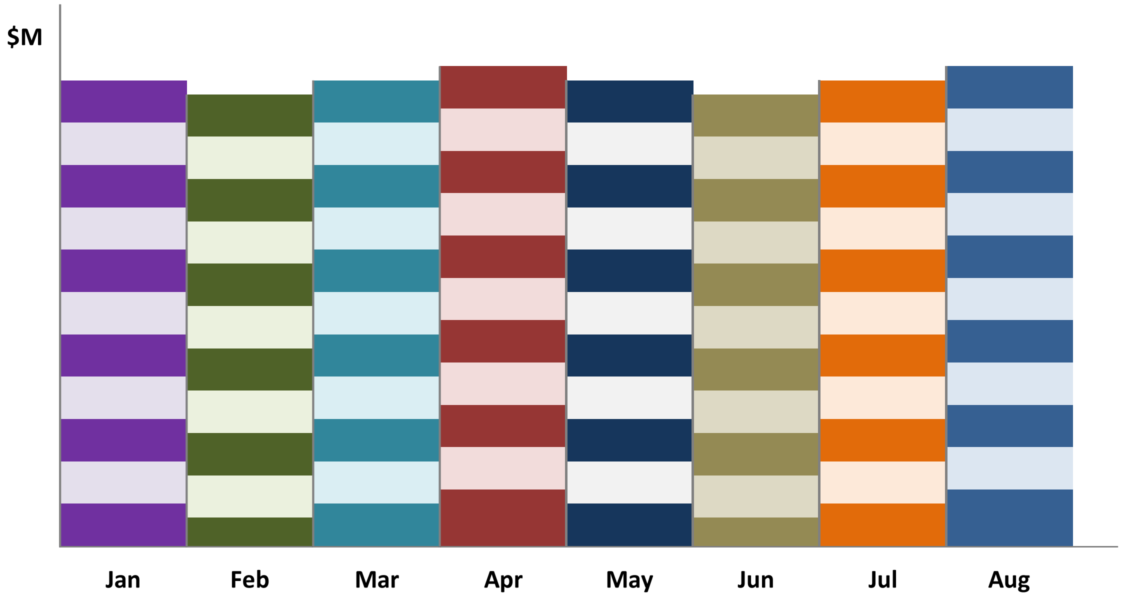

Figure 7-25: The Human Visual System Processes Edges Imperfectly

To avoid edge effects when designing graphics, it suffices to put enough neutral (e.g., white) space between visual elements so that edge effects are no longer noticeable.

7.5.3 Chromostereopsis

Chromostereopsis creates an apparent difference in visual depth between two colours in a visual field [6]. It is caused by the unequal refraction of different wavelengths of light as they pass from the air and across the cornea and lens surfaces of the eye. The result is that when viewers look at a flat visual display, different colours will appear to be at slightly different distances from the eye. This effect can cause cognitive dissonance and be disconcerting to the viewer.

Chromostereopsis has a stronger effect with certain colour combinations. For instance, some bad colour matches are:

- Red and blue
- Red and green
- Blue and dark grey
- Red and grey

Figure 7-26 illustrates chromostereopsis, with a red-blue colour scheme. The human visual system perceives the red bars to be slightly closer than the blue bars, even though they are exactly the same distance away from the viewer.

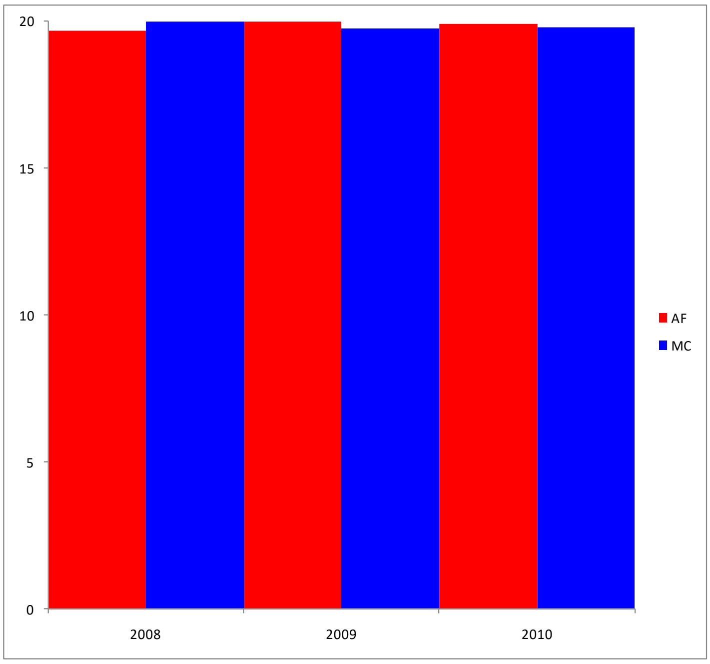

Figure 7-26: Chromostereopsis Creates Apparent Depth

To avoid the effect of chromostereopsis in data displays, it is sufficient to outline the display elements either in black or white, or to put enough neutral (e.g., white) space between visual elements so that chromostereopsis is no longer noticeable. Alternatively, the graphic designer can increase the contrast between adjacent colours, or use a different colour scheme.

7.6 SUMMARY

The effectiveness of data displays depends on how well it matches the preattentively acquired message to the information content that the analyst wishes to convey. Good data display design accounts for the key human factors that affect a viewer's ability to visually analyse the information they contain. Designers of data displays should

1. Assign preattentive visual attributes to the most important data dimensions.
2. Use attentive attributes sparingly.
3. Respect the limits of human visual attention, including change blindness, working memory, and visual distraction.
4. Respect the physiological limitations of human visual processing. Important elements here include lateral inhibition, edge effects, and chromostereopsis.

7.7 REFERENCES AND FURTHER READING

[1] Anscombe, F. J., 1973, "Graphs in Statistical Analysis," American Statistician 27 (February): 17–21.

[2] Bertin, Jacques, 1977, *La graphique et le traitement graphique de l'information*, Paris: Flammarion. (Translation 1981, *Graphics and graphic information-processing*, by William J. Berg and Paul Scott).

[3] Cleveland, William, S., Marylyn E. McGill and Robert McGill, 1994, "The Shape Parameter of a

Two-Variable Graph," *Journal of the American Statistical Association* 83 (402): 289-300.

[4] Cowan, N., Working Memory Underpins Cognitive Development, Learning, and Education. *Educ Psychol Rev*. 2014 June 1; 26(2): 197–223.

[5] Dretske, Fred, 2004, "Change Blindness," *Philosophical Studies* 120 (1-3): 1-18.

[6] Faubert, Jocelyn, 1994, "Seeing depth in colour: More than just what meets the eyes," *Vision Research* 34 (9): 1165–86.

[7] Few, Stephen, 2004, *Tapping the Power of Visual Perception*, Perceptual Edge.

[8] Healey, Christopher G. (2007). *Perception in Visualization*., https://www.csc2.ncsu.edu/faculty/healey/PP/, retrieved 5 May 2022.

[9] Healey, Christopher G. and James T. Enns, 2011, "Attention and Visual Memory in Visualization and Computer Graphics," *IEEE Transactions on Visualization and Computer Graphics* 18(7):1170-88, July.

[10] Kaufman, E.L.; Lord, M.W.; Reese, T.W. & Volkmann, J. (1949). "The discrimination of visual number." *The American Journal of Psychology* 62 (4): 498–525.

[11] Kitaoka, Akiyoshi, Baingio Pinna, Gavin Brelstaff, 2004, "Contrast polarities determine the direction of Café Wall tilts," *Perception*, 33: 11-20.

[12] National Eye Institute, *Color Blindness*, https://www.nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness, accessed 5 May 2022.

[13] Saltzman, I.J. & Garner, W.R. (1948). "Reaction time as a measure of span of attention". *The Journal of Psychology*. 25 (2): 227–241.

[14] Sanford, Ward E. and David L. Selnick, 2013, "Estimation of Evapotranspiration across the Conterminous United States Using a Regression with Climate and Land-Cover Data," *Journal of the American Water Resources Association* 49(1), February, pp. 217–230.

[15] Simons, Daniel and Ronald Rensink, 2005, "Change Blindness: Past, Present, and Future," *Trends in Cognitive Sciences* 9: 16–20.

[16] Summers, Scott H., 2010, "Preattentive Attributes in Visualization Design: Enhancing Combat Identification," Chapter 8 in Andrews, Dee H., Robert P. Herz, Mark B. Wolf, *Human Factors Issues in Combat Identification*, CRC Press.

[17] Ware, Colin, 2002, Information Visualization: *Perception for Design (Interactive Technologies)*, 3rd ed., Morgan Kaufmann.

[18] Wiggins, Rhonda, 1977, "Statistic of the Month: Age Structure of College Enrollment," *American Education* 13:7, 34, Aug/Sep.

[19] Yantis, Steven, 2014, *Sensation and perception*, Worth Publishers.

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**	
		START DATE	**END DATE**
xx-01-2023	Final	Sept 2019	Nov 2022

4. TITLE AND SUBTITLE

Human Factors in Data Visualization, Chapter 7 of NATO SAS-124 Research and Technology Group (RTG), Visualization Design for Communicating Defence Investment Uncertainty and Risk

5a. CONTRACT NUMBER	**5b. GRANT NUMBER**	**5c. PROGRAM ELEMENT NUMBER**
HQ0034-19-D-0001		
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**
C7229		

6. AUTHOR(S)

Thompson, Paul, M.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

IDA Systems and Analyses Center
730 East Glebe Road
Alexandria, Virginia 22305

8. PERFORMING ORGANIZATION REPORT NUMBER

IDA Document NS D-33335
Log #: H 22-000502

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

IDA Systems and Analyses Center
730 East Glebe Road
Alexandria, Virginia 22305

10. SPONSOR/MONITOR'S ACRONYM(S)

IDA/SAC

11. SPONSOR/MONITOR'S REPORT NUMBER

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This document focuses on the role that human factors play in visual data analysis. It treats the basic features and limitations of human sensory and cognitive systems that affect visual data processing. It explains how to leverage these characteristics to design effective data displays, and thereby better communicate results to decision makers or perform exploratory data analysis.

15. SUBJECT TERMS

Data visualization; human factors; cognitive psychology; exploratory data analysis

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		
U	U	U	U	26

19a. NAME OF RESPONSIBLE PERSON	**19b. PHONE NUMBER**
Thompson, Paul, M.	(703) 575-4661