Designing for Understanding: The Cognitive and Historical Foundations of

Successful Data Visualization

Ella Darby

Department of English, Georgia Southern University

ENGL 3235: Information Design

Dr. Joanna Schreiber

Abstract

Data visualization is a product of the relationship between the need for more accessible forms of technical information and the human ability to process visual information much quicker than complex text. As early as the seventeenth century, we began designing categorical data visually, with intentions of opening the door of accessibility of technical information to lay audiences. Modern concepts of encoding methods and accessible design practices provide a foundation for discussions about how the success of specific information design methods is dependent on human processes of comprehending information through comparison and metaphor. Through their straightforward design and ability to create comparative environments for categorized data, bar graphs have become nearly a universally understood form of information design. By examining the history of bar charts and their success through modern concepts of encoding methods, human comparison and metaphor, and accessibility, I aim to explain why bar graphs are so successful today and imply the importance of prioritizing accessible, comprehensible information.

Keywords: information design, historical data visualization, bar graphs, schema theory, conceptual metaphor, accessible design

Table of Contents

Designing for Understanding: The Cognitive and Historical Foundations of Successful Data Visualization

In the digital age, we can represent big datasets in accessible and interactive formats that allow diverse audiences to easily extract information and insights. Additionally, the growing influence of modern tools such as artificial intelligence, virtual reality, analytics suites, and other data discovery tools greatly contributes to ongoing developments in data visualization ("Brief History," 2023). Our extensive toolbox of data visualization techniques has evolved from centuries-old, pioneering iterations of accessible information design that mirror natural human cognitive processes, such as pattern recognition. By encoding technical textual and numerical information through text and shape size, grouping, color, placement, and more techniques, we create signals for relationships between the information and datasets presented. Therefore, the reliable problem-solving quality of effective data visualization lies in its ability to reflect natural human cognition by producing visual, comparative environments.

Foundational cognitive science and psychology studies established that humans have the innate ability to learn through comparison. More specifically, Jean Piaget's schema theory provides that we use preexisting *schemas*, organized knowledge units based on past experiences, to comprehend new information (Nickerson, 2024). Therefore, presenting information in an environment in which users can easily categorize it and draw comparisons opens the doors of accessibility and increased comprehension. As data visualization draws upon visual metaphors, it creates this comparative setting for complex information. Furthermore, visual elements are comprehended much more quickly than complex text. The ability of effective data visualization to harmonize these concepts makes it essential to creating informational content that is accessible to various audiences and understood without a specialized understanding of the topic.

Historic examples from as early as the 17th century provide concrete evidence of how we have been opening the door of accessibility through data visualization for centuries. Furthermore, examining evolutions of information design shows how developments in data visualization result from specific social needs of the time in which they occur. One of the most recognizable methods of data visualization is the bar graph. Although some regard the bar graph as simply boring, yet effective, I argue that this popular graph is so much more. In addition to their simplistic and straightforward design, bar graphs' encoding methods effortlessly harmonize with processes of human cognition. The comparative environment of bar graphs is produced by their use of the universally understood "more is up" metaphor and the schemata associated with related visual metaphors. I provide a discussion of how bar graphs

gained and maintain popularity due to their ability to create comparative and accessible environments for complex, categorized data, that effortlessly communicate with innate human cognitive processes. Therefore, my paper aims to use this discussion of bar graphs as a comprehendible explanation of how data visualization succeeds by reflecting human cognition and how we can use this understanding of data visualization to harness it as an accessibility tool.

The historical need for accessible visual information

Although developments in data visualization are ongoing, they can be traced back to the needs of ancient civilizations. This productivity of data visualization is apparent in early recordings of astronomy, cartography, and agricultural data, including the numerical data represented in the cultures of the ancient Sumerians and Mesoamericans ("Charts and Graphs"). Major historical developments in data visualization have mirrored influxes in mankind's discovery and creation of complex concepts. This correlation is exemplified by the flourishing of data visualization developments around the Age of Enlightenment in the 18th century and again in the Industrial Revolution. Through effective visualization, technical information is made accessible to non-specialized audiences. Therefore, accessible data visualization can influence not only the public's understanding, but also action and decision-making efforts of individuals, institutions, and communities.

Defining data visualization as an accessibility tool

Author of *The Functional Art* (2013) and *How Charts Lie* (2019) Alberto Cairo expresses that "a good graphic should let you visualize trends and patterns without having to read all the numbers" (p. 59). He expresses the purpose of data visualization that was present in even its earliest versions, the goal of accessible information. The "good" quality of visual information that Cairo (2019) speaks of is dependent not only on the reliability and accuracy of the information itself, but also on how this information is presented, or its method of *encoding* (Cairo, p. 79).

Information design uses visual elements, such as color, size, shape, grouping, position, and more, to encode numerical and textual information. These encoding methods act as visual signals that represent relationships between numerical data, so that these relationships don't have to be spelled out within the graph with text or verbally explained. Furthermore, Knaflic (2015) expands on the types of signals that different methods of encoding provide, as she states that users "tend to associate quantitative values" with some graphic attributes, such as size, length, and height, while other attributes, such as color, are

associated with relative value (p. 100). Proper encoding is the key to Cairo's (2019) "good" graphics. Choosing the right methods of encoding is how we understand whether a graph expresses a timeline, comparison, growth/decline, or something else.

Encoding information through comparison and classification

As humans, we depend on our ability to categorize concepts so that we can constantly digest new information and store memories about experiences (Seger & Miller, 2010). This structure is exemplified by our evolutionary history: we had to establish categories for prey/non-threats and predators/threats so that our brains could trigger the right reactions to these situations (Branan, 2010). Scientific studies of visual processing provide that processes of categorization are similar in sighted people and in individuals who were born blind, and that over six major areas of the brain are used during classification in addition to the visual cortex (Branan, 2010; Seger & Miller, 2010). This implies that categorization processes are innate, rather than dependent on or produced from visual inputs. Furthermore, studies of humans' visual processing have established that we "interpret visuals 60,000 times faster than text" ("Visual Journey," 2023). Therefore, effective data visualization allows us to understand numerical values, recognize outliers and patterns, and draw comparisons much quicker than complex textual information.

Understanding encoding through schema theory

The relationships signaled by methods of encoding are understood due to our ability to produce and store *schemas*. Piaget's schema theory suggests that humans comprehend new information by assimilating it into preexisting classifications and interpretations or schemas (Nickerson, 2024). We classify new information through comparison by looking for key characteristics that match these existing mental categories (Nickerson, 2024). Humans are wired to draw conclusions through comparison and classification, which makes properly encoded visual information more approachable and digestible than complex textual information. Additionally, Nickerson (2024) provides that if new information does not fit in an existing schema, it may be "comprehended incorrectly or not at all."

Clear patterns can demonstrate positive or negative outcomes by visualizing encoded increases and declines in datasets through visual metaphors for increase and decreasing values. In successful data visualization, cognitive synthesis between purely numerical datasets and visual design encoding that implies comparison produces a narrative environment. Nickerson (2024) explains, "For learners to process information effectively, something needs to activate their existing schemas related to the new

content." By using size, color, grouping, and other encoding methods, data visualization relies on this innate nature of humans to understand concepts through categorization and comparison.

Linguistic and conceptual metaphors

In schema theory, metaphors are fundamental tools for mapping comparisons and drawing conclusions between new information and preexisting schemata. Metaphors are useful for more than illustrating tangible, educational concepts; they are also used unconsciously in our thinking patterns (Hendricks, 2015). We consistently utilize linguistic and conceptual metaphors in our everyday language. Examples of phrases based on conceptual metaphors include "time is money" and "life is a journey." Based on the connections these metaphors make in our minds, we expand our language to create more abstract, yet systematic, references. For example, the metaphor "time is money" has led us to use phrases like "spend time" and the metaphor "life is a journey" influences our use of phrases like "You've come a long way," in reference to our lives.

Conceptual metaphors are based on a relationship between two factors: the *source domain* and *target domain*. Cognitive psychologists define the source domain as "the conceptual domain from which we draw the metaphorical expressions required to understand another conceptual domain" and the target domain as "the conceptual domain that is interpreted in this way" (Nordquist, 2019). In data visualization, the source domain is the encoding method perceived visually by users, and the target domain is the *relationship* signaled by the encoding method, or the technical information that isn't directly spelled out in a graph.

Graphical and visual metaphors

Risch (2008) provides that visual metaphors presented through data visualization reflect the same systematic process as linguistic metaphors. That is, that they "map the characteristics of some well-understood source domain to a more poorly understood target domain so as to render aspects of the target understandable in terms of the source" (Risch, 2008, p. 1). Risch (2008) relates the success of information graphics to image schema theory, which holds that early experiences and interactions with the world establish structural, spatial patterns that link perception with cognition. Risch (2008) states that in information design, the source domain is a "visuospatial pattern," while the target domain is the information presented in the graphic (p. 2).
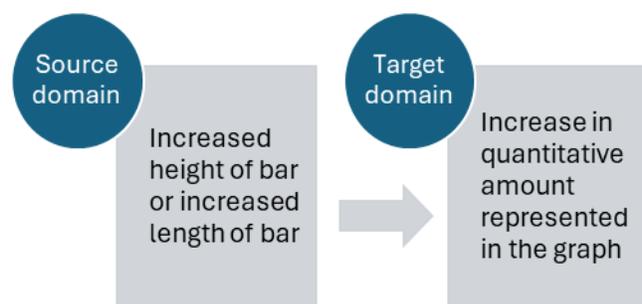
From an early age, and in most cultures, we observe increases and decreases as associated with height and length, quickly establishing the metaphors "more is up" and "less is down." As we add blocks to a vertical tower, or a horizontal line, the amount of blocks increases, and if we subtract them in either circumstance, we observe a decrease. When we get older, we continuously reproduce this metaphor linguistically through phrases such as "prices have gone up," "she has grown up," "the temperature went down," and "graduation rates have gone down." These metaphors are repeated both linguistically and visually every day of our lives, so we quickly understand this signal of an increase/decrease. This metaphor is the foundation of the bar graph, a nearly universally understood method of data visualization. In the next section of this paper, I examine the effect and success of bar graphs as they reproduce this metaphor.

Applying comparison and metaphor to a familiar medium: The bar chart

One of the most common, and universally understood, methods of data visualization is the bar graph. Although bar charts can't be the answer to every data visualization case, they especially demonstrate the relationship between data visualization and our innate ability to learn through categorization. Bar charts depend on height and length to visually encode numerical and statistical data. Therefore, we understand bar charts through the spatial metaphors and schemata we have observed in visual and linguistic environments that imply "more is up" and "less is down." In bar charts, the source domain is an increase in height or length and the target domain is the concept of growth or an increase in the quantity of something. Figure 1 demonstrates the relationship between the source and target domain of the essential metaphor of bar charts.

Figure 1

Graphical metaphor in bar charts

Note. This chart represents the relationship between the source domain and target domain of the "more is up" metaphor employed in bar charts.

Early bar charts and graphs were instrumental in enhancing the accessibility of technical information. Since their earliest innovations, bar graphs provided a setting for information to be presented in a way that engages the human abilities to digest visual information more clearly and quickly than complex text

and understand new concepts through comparison. Today, bar charts are still credited as one of the most reliable chart types as they are "less of a learning curve" for audiences (Knaflic, 2015, p. 56). The essence of the bar chart hasn't changed much since the nineteenth century. Bar charts are a reliable, successful method of data visualization that prime users to "make sense of new experiences through familiar patterns of thinking" through their encoding methods (McLeod, 2024). Therefore, in the following sections of this paper, I will examine the accessibility and productivity of data visualization through a commonly understood and historically enduring medium: the bar graph.

The approachable bar chart

Since its invention around the 18th century, the bar graph has consistently provided a clear, comparative environment for categorical technical information. As they primarily encode information through height and length, bar charts allow users to make quantitative comparisons rather quickly. Through accurately proportioned bar length and height, users can draw conclusions about increases and decreases in categorical data by assimilating the source domain information, the visual height/length of a bar, with schemata about height/length. After this assimilation, users reach the target domain—the comparative technical information omitted from, yet encrypted in, the graph—or the value and narrative that the height/length expresses in the graph's context. By examining a few pioneering bar charts, I provide the problems, data, and environment that produced the essential foundation of bar charts: the reflection of innate cognitive processing.

Early iterations of bar charts and graphs

Figure 2
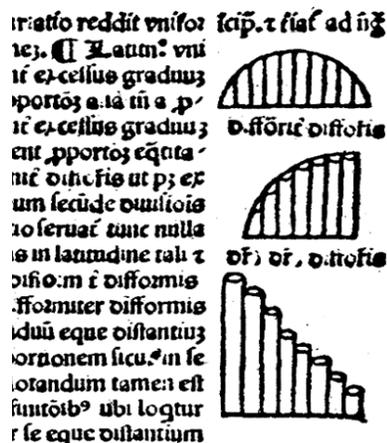
Many sources credit Scottish engineer William Playfair with inventing the bar chart. However, Nicole Oresme, a French philosopher of the 14th century, used somewhat of a bar chart to plot velocity titled *Tractatus de latitudinibus formarum* (Latitude of Forms). Oresme's charts were not based on any data, so they can't be categorized as data graphs like the ones that followed hundreds of years later (Friendly, 2021). However, his use of bars was driven by a unique problem: the need to visually demonstrate a comparative, yet simple environment. Oresme aimed to compare the "velocities of time of an object moving with constant acceleration … to one that accelerated at

a constant rate" by plotting moments of time, or longitudes, along a horizontal line and velocity across the vertical axis ("Nicole Oresme"). Oresme encoded information about the velocity of an object at multiple moments in time through the length of each bar. Therefore, Oresme's innovative charts are an early demonstration of the problem-solving, comparative environment that bar charts provide today.

Joseph Priestley's charts of history

Additional early examples of the use of bars before Playfair are English chemist Joseph Priestley's *Chart of Biography* (1765) and *A New Chart of History* (1769), which plotted horizontal bars across timeline graphs. He examined that these timelines were successful in presenting *synchronisms* or events that took place at the same time. Priestley's charts brought synchronisms into comparison with one another by plotting historical timelines as datasets through bars, therefore, allowing users to visualize the presence of one historical figure in tandem with others. His *Chart of Biography* used horizontal bars along a timeline to encode lifetimes of figures in comparison with others. It provides a comparative narrative for the lives of history's great thinkers, the first five rows of the chart, and its great warriors and rulers represented in the last row of the chart.
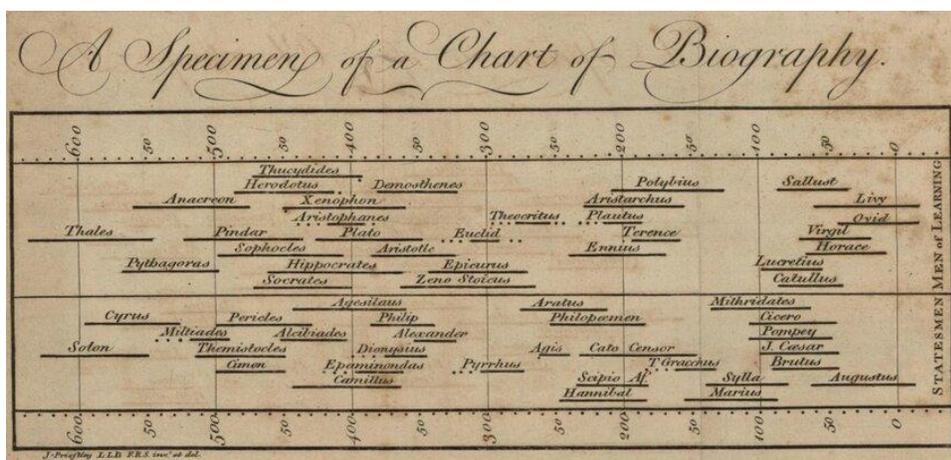
Figure 3

Joseph Priestley's Chart of Biography (1765)
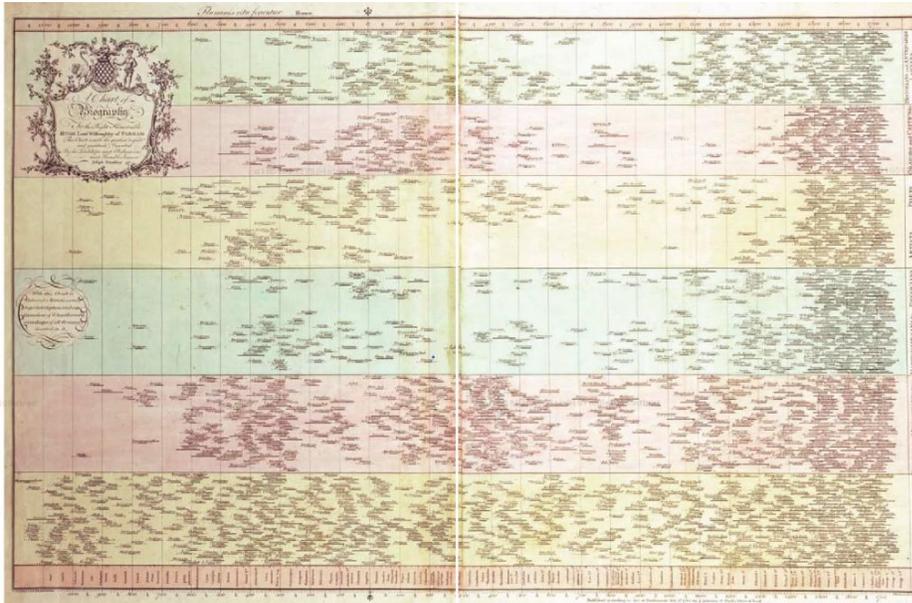
Note. This figure shows Joseph Priestley's *Chart of Biography*, which included the lifespans, denoted by a horizontal line or bar, of over 2000 historical figures over the span of 3000 years. The figures were categorized in relation to their field, with the top five rows including notable figures in the sciences and arts, while the bottom row included history's great warriors and leaders.

Figure 4

*A closer look at Priestley's Chart of Biography (1765)*

Note. Figure 4 provides a closer look at Priestley's *Chart of Biography*. It is a compressed version of the original chart, and only provides two rows. These rows are titled "Men of Learning" and "Statesmen." In this image, it is easier to see Priestley's use of bars along a timeline format and how his graph expressed complex, comparative information.

Four years later, *A New Chart of History*, illustrated in Figure 5, introduced color-coded horizontal bars denoting the presence of major civilizations and empires throughout history. Both of these charts provide productive, visually comparative environments that allow users to quickly point out outliers, like the great wave of notable scientists and thinkers that crests the 18th century during the Age of Enlightenment visible in the *Chart of Biography*. Priestley's bar charts were vastly detailed yet include very minimal textual information. Priestley's humanistic approach to both graphs expresses how data visualization acts as a bridge of accessibility between specialized, technical topics and the common public that lacks an understanding of such concepts. His graphs demonstrate how data visualization assists the user in a way that textual data cannot.

Figure 5

Priestley's New Chart of History (1769)
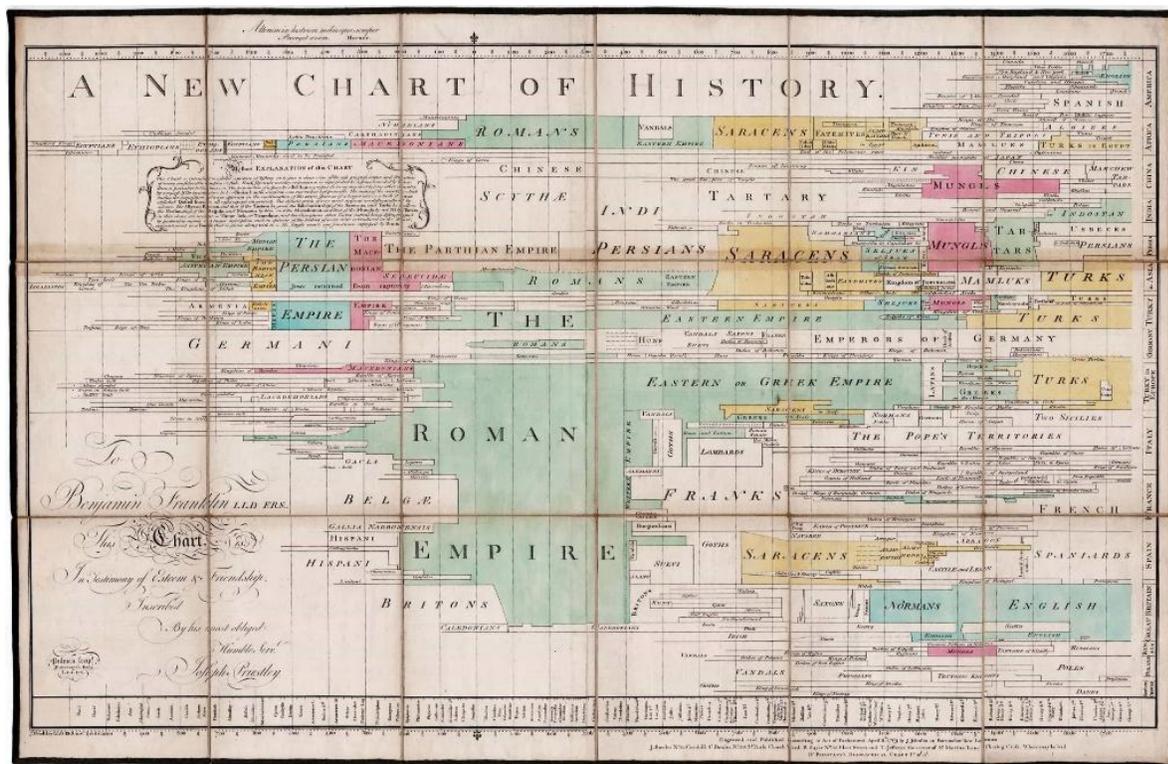
### William Playfair's contribution to data visualization

About 20 years after Priestley's graphs, another problem arose that required visualizing numerical data comparatively. Scottish engineer and political economist William Playfair encountered a unique case in his economic datasets. Playfair published *The Commercial and Political Atlas* in 1786, which included 44

charts and graphs about European populations, revenues, imports and exports, and prices and balances. The *Atlas* introduced Playfair's innovative use of bars which led to him being named the "Father of Statistical Graphs." What made Playfair's bars so revolutionary was the unique problem that they solved. While working on the *Atlas*, Playfair encountered missing time records for trade between Scotland and England in 1780-1781.

He was unable to utilize line charts and examine trade trends over a time series with missing time records. However, he refused to omit this information about Scotland's imports and exports, which led to his invention of using side-by-side bars to compare the data (Friendly, 2021). This pioneering chart, provided in Figure 6, took the form of horizontal bars that provide comparisons between Scotland's imports and exports with each of its trading partners (Friendly, 2021). Playfair expressed that appealing to the eye was essential when presenting "proportion and magnitude," and his data visualization methods allowed for easier, narrative understanding of mathematical and numerical statements. In Playfair's horizontal bar chart, the source domain is the length of each bar, while the target domain is the value of each bar in comparison with the values of others in the graph.

Figure 6

Priestley's horizontal bar graph
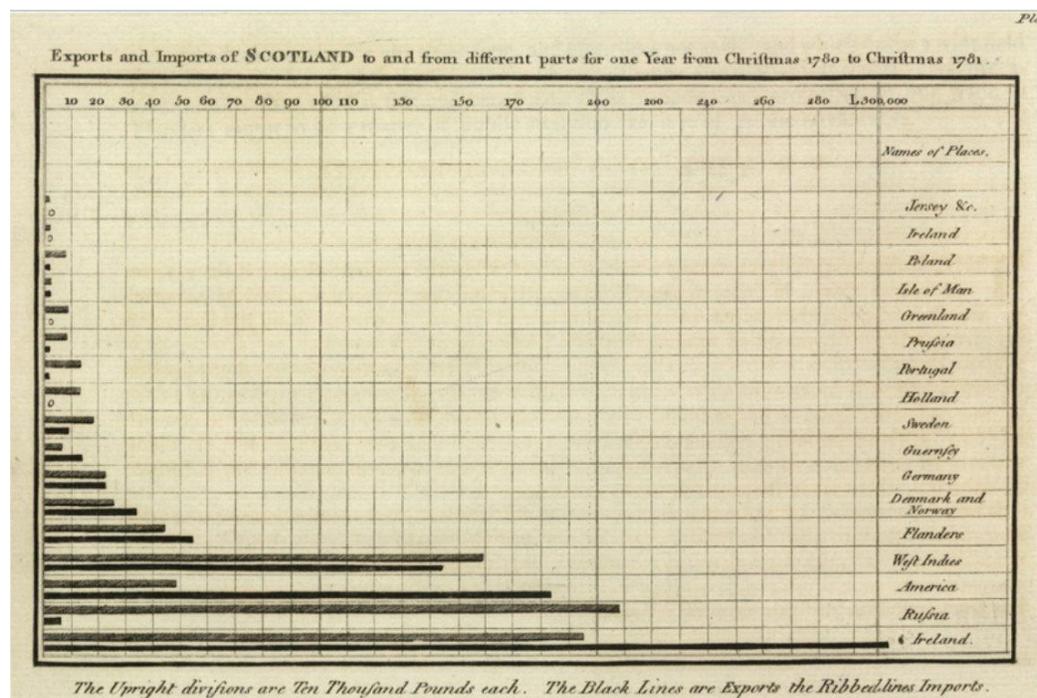
Like Priestley's timeline graphs with bars, and modern efforts toward successful information design, Playfair's *Atlas* aimed to illustrate a comparative and relative environment in which the comparison of categorical datasets tells a story. He expressed that as the knowledge of mankind increased, methods of abbreviating the information became more desirable (Giaimo, 2016). Additionally, Playfair expressed early efforts toward accessible information through data visualization as he acknowledged that reading tables still required a specialized understanding (Giaimo, 2016). He strived to convey technical information without requiring that the audience study the particulars.

Playfair's next project, 1801 publication *The Statistical Breviary*, continued his efforts through more line and bar graphs, as well as the introduction of the pie chart. The introduction to this book includes his statement: "No study is less alluring or more dry and tedious than statistics, unless the mind and imagination are set to work" (Playfair, 1801). The *Breviary* also provides Playfair's belief that learning depends on both how information is *encoded*, or represented in a graph, and how it is *comprehended* by the user (Friendly, 2021, p.96). Playfair (1798) defines visual representations of information as "giving form and shape to a number of separate ideas, which are otherwise abstract and unconnected." His publication was one of the first to explicitly mention an understanding that visual processing is important to comprehending technical data.

This work includes a stacked bar chart, illustrated in Figure 7, which encodes information through color and height in addition to axis labels, data labels, and categorical labels for each decade illustrated. Playfair (1798) expressed apprehension about his somewhat radical publications including a description of why they may be critiqued: "because geometrical measurement has not any relation to money or to time." During his lifetime, Playfair's accomplishments were "often ignored and sometimes denigrated" since they provided visualizations of intangible data (Friendly, 2021, p. 119).

Figure 7

Playfair's vertical stacked bar chart
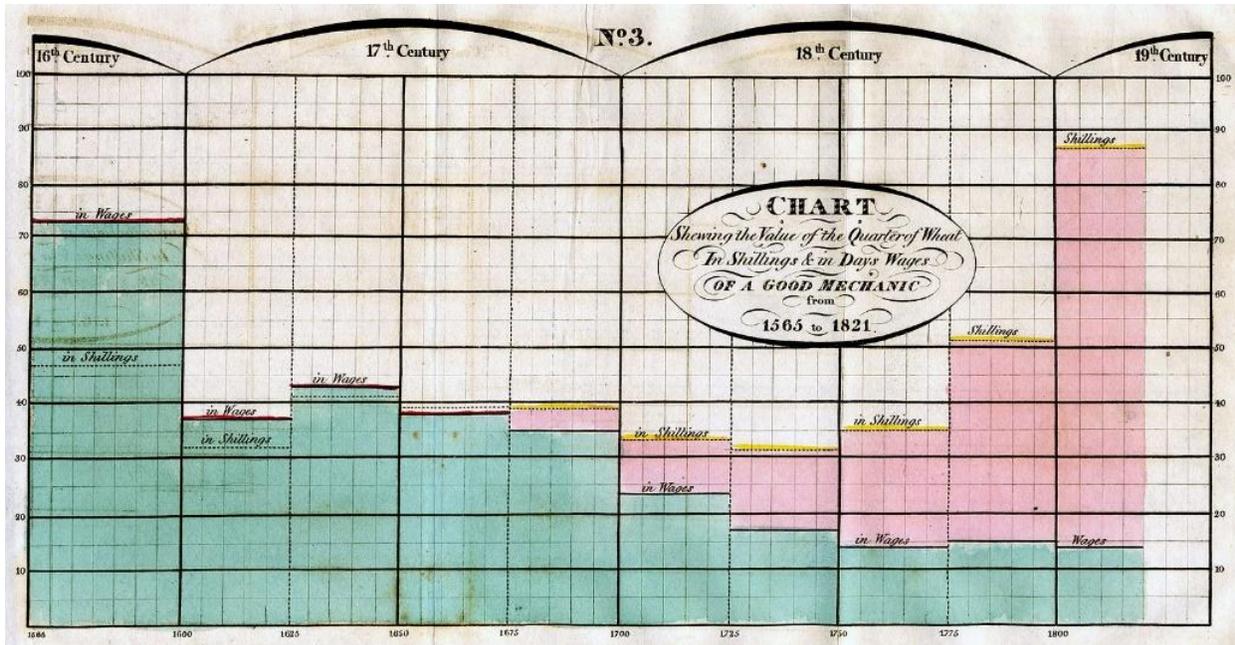
Note. William Playfair's 1801 stacked bar chart shows both shillings and wage earnings of mechanics from 1565 to 1821. He was the first to encode statistical information through color.

Evaluating Playfair's methods today

Although Playfair's bar charts wouldn't pass as entirely accessible design in modern evaluations, his graphs got the encoding fundamentals right. Both Playfair's horizontal and vertical stacked bar charts encode numerical data through length or height. The comparative environment of the bars allows users to draw conclusions about the encoded data as it compares to other categories in the dataset. Furthermore, Playfair's thoughtful, user-centered design approaches also influence the success of his charts. Playfair was skilled at utilizing multiple encoding methods such as color, axis labels, gridlines, and scaffolding to produce effective line, bar, and pie charts. These design components would later become important pillars of user-centered design (Knaflic, 2015; Cairo, 2019).

Therefore, in understanding the productivity of bar charts, it is essential to acknowledge the feats that Playfair's pioneering charts accomplished. However, it is equally as important to evaluate his charts through modern lenses for effective information design to truly understand how we can use data visualization as an accessibility tool. In this section of the paper, I provide some notes on how I would redesign Playfair's first bar chart as well as a visual redesign.

Redesigning Playfair's first bar chart

Knaflic (2015) provides that horizontal bar charts are more accessible than vertical bars as their left-to-right reading structure reflects the visual processing patterns of most users: "starting at top left and making z's with our eyes across the screen or page" (p. 61). While Playfair did format his information in left-aligned horizontal bars, he positioned the "Name of Places" column on the right side of the graph. This makes it a bit difficult to trace the place name to the bars with less data such as those toward the top of the graph. In my redesign of the graph, I left-aligned the place names, bars, graph title, and note. This allows users to seamlessly follow the category to the dataset and utilizes the left-to-right, z-patterned reading structure that Knaflic (2015) explains.

Figure 8

A modern redesign of Playfair's horizontal bar chart

Scotland's imports and exports from Christmas 1780 to Christmas 1781
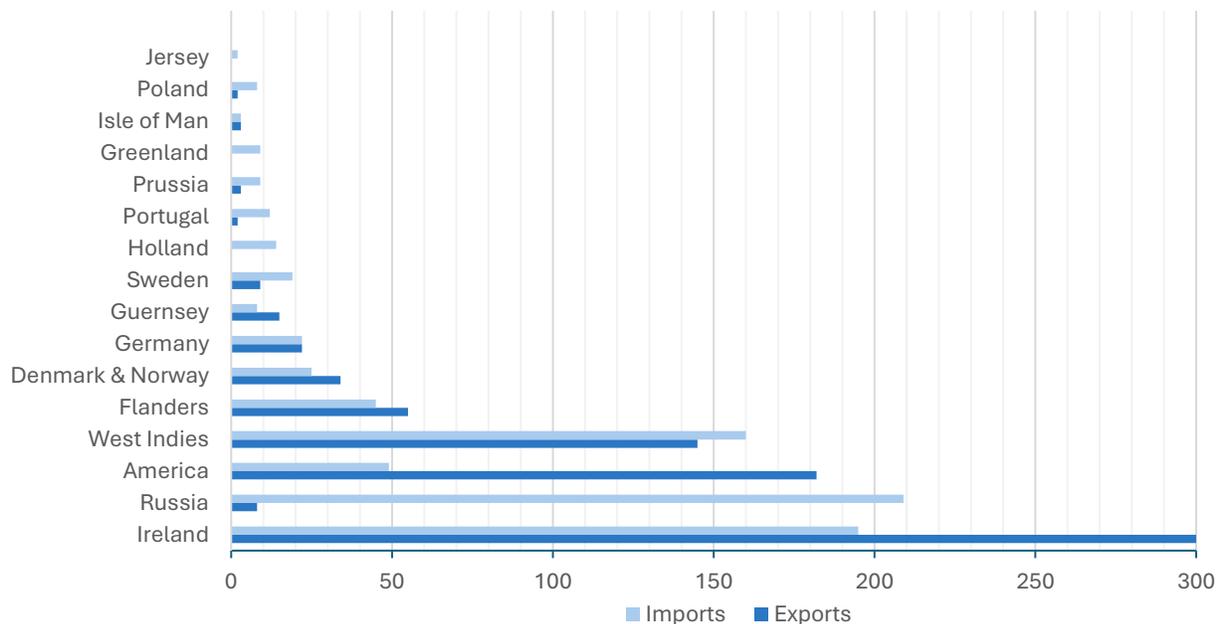
Note: X-axis values are increments of 50,000 denoting the scale: 0-300,000. All values are estimates based

Furthermore, Playfair's chart includes gridlines to demonstrate data values between the increments included along the x-axis. While these gridlines do aid users in understanding the individual values of imports and exports, they clutter up the graph and reduce legibility. In my redesign, I included both major and minor gridlines. The major gridlines, increments of 50, are darker than the minor gridlines,

increments of 10. This design creates a visual hierarchy that allows for a generalized understanding of values to be made more quickly than an understanding of exact values, while still allowing users to draw estimates of those exact values if they choose to. I also employed a graded color scale of medium blues and grays to emphasize and differentiate different components of the graph. Lastly, I included a note about the x-axis values scale, as it demonstrates a scale of 0-300,000 rather than 0-300.

Conclusion: The Role of Data Visualization in Comprehension and Accessibility

Although they look pretty different from the bar charts we see every day, Oresme's velocity bars, Priestley's comprehensive histories, and Playfair's economic data graphs represent the problem-solving abilities of data visualization that interacts with natural human cognitive processing. The goal of data visualization has always been to strike down barriers to accessing and understanding specialized information. By encoding information through recognizable visual representations of conceptual metaphors, data visualization succeeds in producing an intuitive, comparative environment for technical information. It successfully achieves these goals through two key cognitive advantages. First, humans are naturally wired to process visual information rapidly, allowing a single graphic to convey what might otherwise require a lengthy report. Second, humans process new information through our understanding of how it compares to and correlates with our preexisting mental categories or schemas. For instance, encoding information through bars employs users' understanding of a metaphor between "taller/longer bar" and "larger data value" that depends on image scheme theory. Ultimately, effective data visualization provides an easily comprehensible, comparative environment for complex data that we can depend on as an accessible tool.

References

A brief history of data visualization: From maps to BI. (2023, July 12). Insightsoftware.
https://insightsoftware.com/blog/a-brief-history-of-data-visualization/

A landmark chart of history by Joseph Priestley. Boston Rare Maps.
https://bostonraremaps.com/inventory/joseph-priestley-chart-of-history/.

A new chart of history. Stanford Libraries. https://exhibits.stanford.edu/dataviz/catalog/tm577xw2724

Branan, N. (2010, January 1). *Are our brains wired for categorization?* Scientific American.
https://www.scientificamerican.com/article/wired-for-categorization/

Cairo, A. (2019). *How charts lie: Getting smarter about visual information*. W. W. Norton and Company.

Charts and Graphs. Library of Congress. https://www.loc.gov/classroom-materials/charts-and-graphs/.

Giaimo, Cara. (2016, June 28). *The Scottish scoundrel who changed how we see data*. Atlas Obscura.
https://www.atlasobscura.com/articles/the-scottish-scoundrel-who-changed-how-we-see-data.

Hendricks, R. (2015, November 11). *Do metaphors make learning a piece of cake?* Learning and the
Brain. https://www.learningandthebrain.com/blog/metaphors-learning/

History of bar charts and graphs. JPowered. https://jpowered.com/graphs-and-charts/bar-chart-
history.html.

Friendly, M. and Wainer, H. (2021). *A history of data visualization and graphic communication*. Harvard
University Press. *ProQuest Ebook Central*,
http://ebookcentral.proquest.com/lib/gasouthern/detail.action?docID=6631224.

Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*, John
Wiley & Sons, Incorporated. *ProQuest Ebook Central*,
http://ebookcentral.proquest.com/lib/gasouthern/detail.action?docID=4187267.

Nicole Oresme: Father of the modern graph or as he would say: the Latitude of Forms.
https://www.homeofbob.com/science/teacherTools/graphing/info.html

Nordquist, R. (2019, September 8). *Understanding conceptual metaphors*. ThoughtCo.

>   https://www.thoughtco.com/what-is-conceptual-metaphor-1689899.

Playfair, W. (1798). *Lineal arithmetic; Applied to shew the progress of the commerce and revenue of*

>   *England during the present century; Which is represented and illustrated by thirty-three copper-*

>   *plate charts.* https://nla.gov.au/nla.cat-vn3191339

Playfair, W. (1801). *The statistical breviary.* London.

>   https://books.google.com/books?id=Y4wBAAAAQAAJ&pg=PA1&source=gbs_selected_pages&ca

>   d=1#v=onepage&q&f=false.

Risch, J. (2008). On the role of metaphor in information visualization. https://arxiv.org/pdf/0809.0884

Seger, C.A. and Miller, E.K. (2010). Category learning in the brain. *Annual Review of Neuroscience, 33*, pp.

>   203-219. doi: https://doi.org/10.1146/annurev.neuro.051508.135546

The chart of biography. UOregon. https://pages.uoregon.edu/infographics/dev/StoryMaps/Biography/.

The visual journey: Tracing the history of data visualization and graphic communication. (2023, October

>   23). Art of Visualization. https://www.artofvisualization.com/blog/history-and-evolution-of-data-

>   visualization