An Investigation of Emphasis Effects in Data Visualisations

A DISSERTATION SUBMITTED TO MANCHESTER METROPOLITAN UNIVERSITY
FOR THE DEGREE OF MASTER OF SCIENCE
IN THE FACULTY OF SCIENCE AND ENGINEERING

September 2022

By
Sophie Bell
Department of Computing and Mathematics

Table of Contents

List of Tables

List of Figures

Abstract

As society becomes more data-driven and data is considered more important (Kennedy and Hill, 2018), data visualisations become increasingly valuable to present large amounts of information to a wide audience. Visualisations leverage the human visual system (Franconeri, Padilla, Shah, Zacks and Hullman, 2021), further leveraging visual perception could lead to more effective and efficient visualisation designs (Kosara, 2016). This research aims to investigate how preattentive processing links to the design of visualisations. By evaluating emphasis effects applied to bar charts and assessing speed and accuracy when interpreting charts. Previous research has investigated some emphasis effects, including highlighting target elements which is used in the current methodology, however little research has focused solely on bar charts. Designing effective visualisations is important as they are used frequently to communicate information. If designed effectively and the messages conveyed are interpreted correctly, this could lead to better outcomes for those targeted. However, when designed ineffectively or in a misleading manner, this could lead to marginalisation of those who are unable to interpret charts. Researching more effective ways of presenting key information in visualisations could lead to more standardised practices and assist those less able to interpret charts.

The current study was conducted with 31 participants recruited via crowdsourcing on Mechanical Turk. The study used a baseline bar chart, a highlighted bar chart and a horizontal line chart. Participants were required to interpret each of the charts and give the value of the target bar. It was expected there would be a difference in the speed and accuracy for those charts with emphasis effects added. However, Mann-Whitney U Tests were conducted which found no significant differences. Although the null hypothesis could not be rejected, some patterns were identified and the research linking visual perception, preattentive processing and visualisation design were reviewed and presented.

Declaration

No part of this project has been submitted in support of an application for any other degree or qualification at this or any other institute of learning. Apart from those parts of the project containing citations to the work of others, this project is my own unaided work. This work has been carried out in accordance with the Manchester Metropolitan University research ethics procedures and has received ethical approval number 45741.

Signed: Sophie Bell

Date: 30/09/2022

Acknowledgements

Firstly, I'd like to thank my supervisor Dr Anthony Kleerekoper for his guidance and help throughout this project.

Secondly, I'd like to thank my managers Tony and Tom, as well as my other colleagues, for their patience, understanding and support throughout my MSc studies.

Finally, as always, I'd like to thank Lucy for her unwavering support and for inspiring me to work harder and be better every day.

Chapter 1: Introduction

1.1 Project Overview

Data visualisations are widely used in academic and non-academic settings to analyse data and communicate information to a variety of audiences (Kennedy and Hill, 2018). Various assumptions have been made in early research for the best way to design visualisations, but there is little empirical evidence on what is the best chart design in any given situation (Kosara, 2016). Various techniques are used with little knowledge and evidence about best practices as early research (Cleveland and McGill, 1984) and recent research (Kosara, 2016) suggests data visualisation needs a scientific foundation. The current research aims to measure the speed and accuracy of participants estimates of bar charts that have been subjected to different emphasis effects with the overall aim to contribute to the existing body of data visualisation research and link with psychology research due to the visual perception and preattentive processing literature.

The present research aims to answer the following research questions:

1. Do readers extract information from a bar faster and more accurately when emphasis effects are used?
2. Is there a difference in speed and accuracy between a standard bar and one with emphasis effects?

1.2 Purpose of the Report

The purpose of this study is to investigate how different emphasis effects on bar charts effect the speed and accuracy of value estimation. The secondary purpose is to contribute to the growing literature on data visualisation design that utilises visual perception and preattentive processing informed designs to improve readability and efficiency.

1.2.1 Aims and Objectives of Research

- Contribute to literature on data visualisation design
- Investigate the use of preattentive processing in visualisation design
- Assess the effectiveness of emphasis effects in bar charts
- Make suggestions based on this preliminary research on future emphasis effects
- Conduct an experiment to evaluate the effectiveness of emphasis effects in bar charts

1.3 Report Structure

This report is split into chapters covering the various stages of the research project. Each chapter will cover a different area, these areas and chapters are summarised below.

- Chapter 2 provides a review of the literature covering numerous areas relevant to data visualisation research, including visual perception, bar chart design and previously presented theories of visualisation design.
- Chapter 3 will cover the data collection process and justification for the method of data collection, using Mechanical Turk.
- Chapter 4 will cover the experimental methodology, describing the methods used and giving examples of questions asked of participants.
- Chapter 5 will present the data analysis techniques used and provide and interpretation of the results.
- Chapter 6 presents a critical evaluation of the presented research.
- Chapter 7 concludes this report, summarising the findings linking this with the Chapter 2 literature review. Limitations from the current research are given, and building from this, suggestions for future work are given.

Chapter 2: Literature Review

2.1 Introduction

There are a wide variety of topics to consider when investigating data visualisation design and the interpretation of data visualisations. These include:

- Visual Perception
- Data Visualisation and Bar Charts
- Graphical Literacy
- Visual Encoding
- Emphasis Effects.

The above will be discussed in detail starting from subsection 2.2, followed by a summary of research relevant to the current project.

2.2 Graphical Literacy

Visualisation literacy or graphical literacy has been defined as a concept generally understood as the ability to confidently create and interpret visual representations of data (Boy, Rensink, Bertini and Fekete, 2014). Okan, Garcia-Retamero, Galesic and Cokely (2012) describe graphical literacy as a skill typically acquired through formal education. Given this, the level of a reader's graphical literacy can be a factor affecting how accurately charts are interpreted. Studies conducted on graphical literacy will be presented, along with the importance of graphical literacy in society and how this relates to the current research.

Okan et al., (2012) have summarised the three stages of graph comprehension. The first is encoding the visual pattern, which includes judging the elements of the chart. For example, the position on a scale (x axis) and length of the bar (y axis) in a bar chart. The second is translating the visual features into conceptual relations, this includes translating different elements. This is key to the current research as this process includes viewing the saliency of charts and interpreting its meaning. An example would be using colour to indicate variation. Finally, the third process involves decoding the graphical concepts and associating them with specific variables and their numerical values (Shah & Carpenter, 1995: cited Okan et al., 2012). The third stage includes identifying information such as the scales, axes and numerical values to relate the information being presented to real world items (data).

Okan et al., (2012) examined the effect of graphical literacy levels on chart comprehension and found participants who scored higher in graphical literacy spend more time looking at key information in the graph (such as y-axis labels) and were more likely to give accurate responses. Graphical literacy could be a factor in the current research as the key metric is accuracy of estimation. Even though the data suggested those with lower literacy levels do not use the axes for interpretation, outcomes may improve by highlighting a bar and removing complexity in the chart (by highlighting the main focus). Therefore, it may be easier to interpret information from the axes. In addition, Okan, Galesic and Garcia-Retamero (2016), conducted further research using eye-tracking methods, they again found that individuals with higher graph literacy spend more time looking at conventional features in charts. They identified individual differences in graph comprehension, linked to reader's level of graphical literacy. This further supports graphical literacy being a key topic in data visualisation research and an important consideration for the study of how bar charts are read as it defines the concept of a user's level of understanding of charts and how they are created (Okan et al., 2016). A user with low visual literacy may struggle to interpret the information being presented in a graph and thus misunderstand the message being communicated to them (Okan et al., 2012). By further researching the link between graphical literacy and visual perception, researchers can help design guidelines to ensure data visualisations can be understood by all the target audience.

It is crucial techniques are developed to provide effective access to electronic resources so the information is more easily interpreted by all, rather than those with higher graphical literacy (Elzer, Carberry and Zukerman, 2011). The importance of creating good, accurate and understandable visualisations becomes more apparent in research into graphical literacy. For example, individuals who struggle to interpret information from graphs and other visual data sets may be at a disadvantage than those who possess this ability (Boyd and Crawford, 2012). Data visualisations become the most appropriate way to address this issue as visualisations are cited to make information more transparent and accessible (Few, 2008; Zambrano and Engelhardt, 2008).

Graphical literacy is important to data visualisation design and has been researched extensively. More recently, the idea of datafication (Mayer-Schonberger and Cukier, 2013) has been discussed. Datafication is the contemporary phenomenon of quantifying aspects of life that did not exist numerically so that it can be tabulated and analysed (Mayer-Schonberger and Cukier, 2013). As datafication within everyday life increases so does the need to understand charts presented for either communication or analysis and those without the skill to interpret this information can find themselves marginalised or disadvantaged. Therefore, as the amount of data or quantified things increase, so does the average skill level for graphical literacy needs to be improved. In order to assure that target and sometimes mass audiences can interpret visualised information. Kennedy and Hill (2018) further suggest that data within society is an ever-growing importance due to an increasing number of decisions that affect the population are being based on data. Hypothetically, potential important decisions or judgements may be made from misinterpreted data (Zarocostas, 2020).

Low graphical literacy and ineffective design can lead many viewers to struggle to understand these otherwise powerful thinking tools (Franconeri, Padilla, Shah, Zack and Hullman, 2021). Public policy visualisations can be counterintuitively designed, leading many viewers to draw a conclusion opposite the one suggested intended (Engel, 2014) or intentionally presented one way to direct towards certain ideals (Dick, 2015). Glazer (2011) has stated it is a crucial skill to be literate in data, being able to extract patterns, find trends and criticise data. Given these assertions in the research, designing charts effectively is important to ensure the right message is being presented, particularly when it is relevant to a population's health/wellbeing.

Graphical literacy is an important social issue for many reasons, firstly because visualisations are a popular method of communication as previously discussed and secondly because those who do not have some graphical literacy can be further marginalised. Some researchers are calling for investigations into graphical literacy and considerations for interpreting graphs when the skills to read them have not been taught. There is a large potential for data visualisation to improve the lives of people who are only recently gaining access to the internet and increasing amounts of data. Improving data visualisation design and considering those with lower graphical literacy could lead to improvements in emergent ICT users' lives (Jena et al., 2021). An example, a banking application allows users to transfer money and view balances, there are also options to view spending patterns and trends within visualisation, those who cannot interpret charts are missing out on the opportunity to further control and manage their own money (Jena et al., 2021).

2.2 Data Visualisation

The definition of data visualisation adopted throughout this research is that proposed by Chen et al., (2013): "a field of study concerned with the transformation of data to visual representations, where the goal is the effective and efficient cognitive processing of data". Interestingly, this definition states the goal of visualisations is to effectively and efficiently process data. The current research will be using different methods of highlighting columns in bar charts to see if the efficiency and effectiveness of interpreting the charts is improved.

Data visualisation is the process of representing data in a graphical or pictorial way in a clear and effective manner (Sadiku, Shadare, Musa and Akujuobi, 2016). There are many cited definitions for data visualisations and numerous synonyms used in the literature. Alongside the definition of a data visualisation, there are also the goals of a visualisation. Knowing the goal of a visualisation is crucial for the evaluation of visualisation (Heer, Bostock and Ogievetsky, 2010). There are multiple theories posited on the differences between goals and definitions of different types of visualisations. For example, information visualisation aims to explore abstract data and create new insights (Heer et al., 2010) whereas knowledge visualisations aim to improve the transfer of knowledge (Heer et al., 2010). Knowledge visualisations align more to the definition used for data visualisation (transfer of knowledge and effective processing of knowledge). In addition to this, the research aims to measure ways to improve the efficiency of information found in bar charts by highlighting columns. Further to this, Lewandowky and Spence (1989) stated graphs can be used in two different ways, the first being to communicate information to an audience and the second to analyse data. Chen, Floridi and Borgo (2013) have argued that defining a visualisation and its goals are important because if not defined it will not be possible to know the effectiveness. If it cannot be determined how effective visualisations are it will be difficult to objectively measure and validate them and this will lead to a lack of development in visualisation research (Chen et al., 2013).

Insight is what visualisations aim to achieve, however, Chen et al., (2013) have described the notion of insight as elusive, difficult to measure and evaluate accurately or validate objectively. Using time in the definition means that visualisation can be measured more objectively, for example, in empirical studies response time can be used as a measure alongside accuracy, this is proposed to be the most fundamental objective by Chen et al., (2013). Whereas 'insight gained' would also need to consider previous knowledge as participants which would make obtaining a fair sample or designing a fair study extremely difficult and difficult to validate. Ware (2004) states one of the major advantages of data visualisation is the amount of information that can be interpreted quickly, when presented well.

The goal of a visualisation is to aid our understanding of data by leveraging the human visual system's highly tuned ability to see patterns, spot trends, and identify outliers (Heer et al., 2010), and when designed effectively visualisations influence the visual system's processing power to allow faster processing of data (Franconeri, Padilla, Shah, Zacks and Hullman, 2021). When designed ineffectively, these displays leave viewers confused and unable to interpret new displays. This further highlights the need to investigate the most effective ways to design bar charts for different circumstances to improve the efficiency of the data interpretation.

2.3 Bar Charts

A bar chart displays quantitative values for different category items. The chart comprises line marks (bars) with the size attribute (length or height) used to represent the quantitative value for each item (Kirk, 2016). Bar charts are used across multiple areas, from academic research to use in the media and are arguably the most ubiquitous visualisation techniques (Zhao, Qu and Sedlmair, 2019). Their usage ranges from casual use for personal data to professional decision makers (Zhao et al., 2019).

Mogull and Stanfield (2015) conducted research on the most common chart type published in Science (2014 edition). Across a randomly selected sample of articles, 91% contained a chart, of those, the most common type (27% of the charts) were bar charts. In another study, Borkin, Vo, Bylinkskii, Isola, Sunkavalli, Oliva and Pfsiter (2013) analysed 2,000 visualisations from around the web and found over 40% of charts published by news outlets were bar charts. In the study conducted by Borkin et al., (2013) it highlights how important an understanding of bar charts is when news outlets use them frequently, an understanding from the public consuming news is required to accurately interpret information being presented.

Research has shown that there are many factors affecting how bar charts are read, as well as the length of the bars when interpreting the information. These include the width of the bar, the groupings of the bar, the overall shape of the bars (though mostly reserved for histograms) and the distance between the bars. There are also a variety of bar chart configurations in the literature including: stacked charts, grouped charts, 3D bar charts, horizontal bar charts and histograms. The current research will focus on vertical bar charts only, as there are many factors to consider for each of these charts and little empirical evidence. Grouping them separately allows for more specific research to be conducted which can lead to separate design guidelines for different use cases.

2.4 Visual Perception

Visual perception is a heavily studied topic in psychology, dominating most perception research (Eysneck and Keane, 2015). Visual perception is a key topic in data visualisation as the aim is to understand how people view information and how it is interpreted within the brain (Ware, 2019). If visual perception is understood, the knowledge can be translated into guidelines for displaying information (Ware, 2019).

There are two concepts in the perception research key to data visualisation design: preattentive attributes and preattentive processing. Preattentive attributes, have been grouped into four categories: form, colour, position and motion (Ware, 2019). Few (2006b) further defined eleven types of visual attributes, one being line length. Few (2006b) states that length of a line and position are the best attributes for encoding quantitative data, this research further confirms Cleveland and McGill's (1984) work stating length of line is second only to position when looking for the most accurate method of comparison, this was confirmed again in research conducted using crowdsourcing methods (Heer and Bostock, 2010). Preattentive processing occurs below the level of consciousness and is tuned to detect a specific set of visual attributes (Few, 2006b). It involves visual features that can be detected by the human visual system without considering a particular region of the image (Healey, Booth and Enns, 1995). How much something is noticed preattentively depends on how different the highlighted element is from the others in the group. For example, it is easier to see a different number in a list of numbers when the number is darker in colour (Fronza, Janes, Sillitti, Succi and Trebeschi, 2013). According to Bartram et al., (2017), colour is preattentively observed, making it particularly useful in conveying qualitative and quantitative information in maps and diagrams. Given this assertion, preattentive processing is the most relevant area to the current research as these processes are leveraged by drawing attention to areas that are processed subconsciously. This will be the aim of adding highlighted bars. Kosara (2016) suggests that information visualisation should be designed by understanding the principles of perception, test how they apply to different data encodings, then build up those encodings to see if the principles apply to visualisation. Further to this Seong et al., (2020) encourage designers to take advantage of the human visual system's ability to do preattentive processing by seeking to encode information preattentively.

Barrera-Leon, Corno and Russis (2022) conducted a literature review with the aim of understanding how concepts related to preattentive processing are used in recent research relating to data visualisation. The literature review found there are two areas preattentive processing and visualisation research fall into, the first is design components. Within this research, researchers investigate how manipulating attributes can achieve reaching different goals such as highlighting a specific subset of data within a chart. The other school of research centres using preattentive processes as a measuring tool such as using the preattentive process to validate design decisions (Barrera-Leon et al., 2022).

2.5 Visual Encoding

Cleveland and McGill (1984) conducted experiments measuring how accurately participants were able to perceive the quantitative information encoded by different cues. This will be referred to as visual encoding which is the process of converting images and visual sensory information to memory stored in the brain (Eysenck and Keane, 2015). Cleveland and McGill (1984) ranked the accuracy of each of these visual encoding options based on participants responses. The results showed that when measuring responses some visual encodings produce more accurate responses than others. The encodings are (in order of accuracy): position, length (aligned on x-axis), slope, angle, area, colour intensity, volume and colour hue. This paper is widely cited in visualisation literature as it is among the first to rank visual cues in relation to accuracy. There are some criticisms to the approach taken by Cleveland and McGill (1984), such as the limited number of participants used which is often cited as a downfall in visualisation research (Kosara, 2016) and the measurement of accuracy. Although the aims were based on the accuracy of participants responses, this is not necessarily the only way to measure if a chart is useful or appropriate in other circumstances. The aim of a chart is not always to present accurate information, otherwise tabulated data could be the most efficient method, it can also be to demonstrate pattern or groupings or communicate a message. Here accuracy becomes less important, and the decreased time taken to reach a conclusion or the message received from the visualisation could be the most important factor.

More recent research has shown there are different types of encodings beyond the visual encoding used for a chart (Kosara, 2022). These are specified encodings which are the visual properties that are used to create the visualisations such as the length of a bar. Observable encodings are visual properties of the visualisation that can be the same or related to the specified encodings, for example, the area or aspect ratio in addition to length of a bar in a bar chart. Observed encodings are those that the reader uses from the visualisation, however, Kosara (2022) states that it is not possible without further research to distinguish between observable and observed encodings.

Kosara (2022) argued that visual encodings are assumed to be read exactly as specified, although this is not always the case. Based on early research, it was assumed that pie charts are read by looking at the angle of the sectors based on the centre (Kosara, 2016). But more recent research has shown this is not the case and arc length or area are the visual cue most likely to be used (Skau and Kosara, 2016). One difficulty with making a distinction between the assumed visual encoding and how the charts are read is that the assumed and used often behave the same. In the pie chart example, a pie slice's area and arc length (the visual cues used) increase linearly with the angle of the line (the assumed use).

The general shape of a bar chart can also be read separately to the values shown by individual bars. For example, a reader may look for a bar that does not follow the same increasing pattern or a bar in the middle that is shorter than its surrounding bars and appearing to break the pattern (Kosara, 2022). In terms of visual encodings, a bar chart would have the following: specified encodings are length, and the observable encodings include length, area, aspect ratio and a derived encoding would be the enveloping shape (such as those shown in histograms) (Kosara, 2022).

Another concept within the visual perception and encoding research is visual salience. Visual salience measures how much an item stands out with respect to neighbouring items. The higher this value, the more visual attention it attracts. This suggests that the effectiveness of a visualisation can be improved by having a desired level of salience (Janicke and Chen, 2010). Visual salience defines which parts of the image stand out and will likely attract a lot of attention, ideally this should coincide with the parts of the dataset that are important to the message (Janicke and Chen, 2010). In a study by Janicke and Chen (2010) they argue that salience should be used to measure the quality of a visualisation as it is more generalisable across different disciplines and chart types and it focuses on measuring how much the target data is highlighted in comparison to other data, guiding the user's attention (Janicke and Chen, 2010).

Preattentive processing in visualisation means a person's gaze moves to a salient object before the person is consciously aware of the object being present. This is relevant to the current research as it is expected that more salient variables would guide a viewer's attention to the highlighted bar and allow a faster and more accurate estimation of value.

It is often argued in literature, that charts should be kept minimal, and the data-ink ratio should be as high as possible (Tufte, 1983, Cleveland and McGill, 1984). This means that most ink on the page should be dedicated to representing data, not other embellishments. However, more recently, Bateman, Mandryk, Gutwin, Genest, McDine and Brooks (2010) conducted research comparing minimal charts to highly graphic charts, measuring memory and readability. The study found no difference in people's interpretation accuracy between minimal and embellished charts, concluding that visual embellishments do not hinder a reader's ability to interpret information. Embellished charts were found to be more memorable then their minimal counterparts (Bateman et al., 2010), which could have implications in some disciplines/media outlets where the message is to persuade a reader.

2.5.1 Emphasis Effects

Emphasis effects are the techniques used to emphasise an element of a visualisation with the aim of drawing attention or indicating importance (Mairena, Dechant, Gutwin and Cockburn, 2020). The goal of emphasis to manipulate the visual features of an important data point to make it visually prominent, such that a viewer's attention is attracted to the point (Hall, Perin, Kusalik, Gutwin and Carpendale, 2016). There is limited guidance and empirical evidence on choosing emphasis effects or how these affect readers' interpretation of charts (Mairena et al., 2020). Further to this, literature has stated it can be difficult to know when designing visualisations how different emphasis effects compare to each other. The present research is important to practitioners as when designers are using emphasis effects, it is important they know the merits of each effect.

There are two types of emphasis effects, time-invariant and time-variant. Time-invariant effects are those that do not change over time, such as changing the colour of a bar. Whereas time-variant do change over time, such as a fade-in effect (Mairena et al., 2020). There are many emphasis effects that have been discussed in research; blur, size, motion, colour, shape, flicker and fade-in (Mairena et al., 2020). The present research will focus solely on time-invariant effects.

2.6 Related Research

Baldonado, Woodruff and Kuchinsky (2000) conducted research on using multiple visualisations in various scenarios, the authors suggest multiple rules when designing multiple visuals. Although the research was designed to consider more than one view, it could be applied to single charts to allow for more better investigation of data by a user. The rule of self-evidence from Baldonado et al., (2000) states that visualisations should use perceptual cues to make relationships along multiple views more apparent to the user. Although in this research there is only a single visualisation, the same rule of using perceptual cues to make data points more apparent to a user applies, in this instance, it is a single bar amongst a group of bars.

Authors have defined highlighting in three layers, the first layer is the traditional view-based visualisation definition, where highlighting acts as the viewing control to attract a user's attention into a portion of the visualisation. The current research aims to contribute further to this by investigating a highlighting technique that is used at the visualisation viewer level. The current research will also contribute further to this work as it is investigating the specific use of highlighting an element of a bar chart with the aim of increasing the efficiency of reading a bar chart which is in line with Liang and Huang's conclusion (2010). Other terms for highlighting include emphasis effects as discussed in Section 2.5.1, cue-based techniques, attention retargeting and focus+context (Waldner, Karimov and Groller, 2017).

The current research is investigating preattentive processing as a design component, examining how designing charts with preattentive processing elements can improve efficiency of chart interpretation. According to Bartram et al., (2017) colour is preattentively observed, this characteristic makes it particularly useful in conveying qualitative and quantitative information in graphic illustrations and diagrams. Colour is the most common preattentive attribute used to improve the comprehension of graphics (Barrera-Leon et al., 2022). Seong et al., (2020) encourage designers to take advantage of the human visual system's ability to do preattentive processing by seeking to encode information pre-attentively visually.

Mairena et al., (2020) carried out two studies investigating three different emphasis effects (colour, size and blur). They noted the difference between these variables, colour and size edit only the target bar, whereas the blur emphasis affects everything but the target bar. The aim of the study was to determine whether different effects are perceived differently, and to provide evidence on how the emphasis effects compare to each other. The study focused on scatter plots only using time until first fixation on the target point, time to click the variable and a subjective rating from participants on the emphasis effects. The results show consistent differences in measures across the multiple emphasis effects Mairena et al., (2020). The study found blur to be the most effective emphasis effect, arguing this is because the entire visual is manipulated, apart from the target bar.

Given this research was conducted on scatter plots and the current is focused on bar charts there are large differences. However, providing empirical evidence on emphasis effects is the aim of the present research and to contribute to visualisation design studies. It could be expected that scatter plots, would be more susceptible to size as an emphasis effect as in any circumstance this would be something that would draw the eye, especially when compared to a standard size. Similarly, blur would arguably have a different effect on a bar chart as bars are beside each other and can be wide depending on the number of categorical variables. The similarity with this study and the previous lies with colour as an emphasis effect. The authors also argued that there is a maximum difference between the colours chosen in a visualisation, whereas there is no upper limit on size. The study compared different magnitudes of size and colour, however, there is no mathematical way of calculating the difference in colours and the research used the same colour but changed the brightness/hue, suggesting that size would create more of an emphasis effect as there is a larger magnitude for difference. The present research will be using colour as an emphasis effect; however, the colour will change completely utilising turquoise, orange and yellow rather than shades of red and blue. The current study is aiming to evaluate the difference between the chosen emphasis effects and a baseline bar chart and measure effectiveness rather than comparing magnitudes of emphasis effects.

Healey et al., (1995) conducted an investigation with the aim of determining whether preattentive processing can help design more useful visualisation tools. The results showed that rapid and accurate estimations were performed when hue and orientation techniques were used to capitalise on preattentive processing. This was measured by asking participants to estimate the value, some of which were modified to harness preattentive cognitive processing. The present research will use a similar methodology to measure responses to different emphasis effects that leverage preattentive processing. Though there are some methodological differences, the present research is conducted online to achieve a wider range of participants.

2.7 Summary

The above literature review has covered topics relevant to data visualisation, including, a definition of data visualisations and bar charts. Visual perception and visual encoding and the relationship they have to data visualisation design have also been discussed in depth. The different emphasis effects that have been used as well as recent research on the effects. The present research aims to build on the above, utilising well researched theories of visual perception and preattentive processing to improve the research conducted on data visualisations and ultimately the future design of data visualisations.

Preattentive processing has been described as an important design component for data visualisations and utilising these can improve the efficiency of charts. This research builds on a suggestion from Kosara (2016) to use visual perception to design charts and is further suggested by Barrera-Leon et al., (2022). By understanding visual encoding from a perception perspective, chart designers will be able to leverage the way information is processed to enable more efficient chart design. Emphasis effects have been studied previously on a variety of chart types. However, the present research will investigate two different emphasis effects, specifically on bar charts, building on emphasis effect research and drawing on knowledge from visual perception research.

Previous research has shown colour to be processed preattentively (Bartram et al., 2017), which is thought to lead to faster processing of information (Janicke and Chen, 2010). Given this, the current research will investigate the usage of a highlighted bar on participants response time and accuracy.

Hypothesis 1: Using a highlighted bar as an emphasis effect will results in faster response times from participants, compared to a standard bar chart.

Hypothesis 2: Using a highlighted bar as an emphasis effect will result in more accurate response times from participants, compared to a standard bar chart.

In other research, visual encoding is discussed in relation to how charts are interpreted (Kosara, 2022). There are a variety of ways a bar chart can be interpreted, none of which are empirically evidenced. The second emphasis effect to be investigated is a horizontal line from the y-axis to the edge of the chart as a bar replacement. Currently, no research has been conducted on this, but based on preattentive processing research. The current study will investigate its performance compared to a standard bar chart. It is thought the y-axis is important when interpreting bar charts (Kosara, 2022, Dick, 2015), therefore by utilising a coloured line to leverage preattentive processing, it could lead to faster interpretation of a value shown on the y-axis.

Hypothesis 3: Using a horizontal line to replace a bar will lead to more accurate responses from participants, compared to a standard bar.

Overall, based on the preattentive research and emphasis effect research cited, it is expected there will be a difference in accuracy and speed of the emphasis effect charts, compared to the standard bar chart.

Graphical literacy is a key topic in visualisation research as it can account for individual differences in chart interpretation and lead to design choices being made to improve readability for all audiences with varying chart knowledge.

Hypothesis 4: Those with higher familiarity with charts will answer more questions correctly.

Chapter 3: Experimental Methodology

3.1 Introduction

The following sections will describe the methodology used in the current research and justify the methods used in relation to other published literature and the aim of the research. Firstly, crowdsourcing will be discussed as a method and justified as a means of data collection. Secondly, the online questionnaire design will be discussed. Finally, the process of data collection will be described in full, including the design of the website and databases for questions and responses.

3.2 Amazon’s Mechanical Turk

A common criticism of visualisation research is the overuse of the student population (Lam, Bertini, Isenberg, Plaisant and Carpendale, 2014). As discussed in Chapter 2, visualisations are seen and require interpretation by most of the population as they are a common method of communication in a variety of areas (Zarcostas, 2020). Therefore, it is important to evaluate visualisations with a wide variety of the population, not just those at university. In response to this, more research uses Amazon's Mechanical Turk (MTurk) as a means of crowdsourcing research responses. MTurk is an open online marketplace for getting work done by others (Buhrmester, Kwang and Gosling, 2011). Buhrmester et al., (2011) have investigated MTurk and its use in research and have found in an analysis of demographic characteristics that MTurk participants are at least as diverse and more representative of non-university populations that those of traditional samples. This is ideal for the current research as the reach of data visualisations and therefore the importance of the design extends beyond the student population. An overreliance on the student population as participants has also been cited as a criticism for other visualisation research (Lam et al., 2014). They also found that worker motivation was intrinsic as well as monetary. The large amount of MTurk users can supplement or replace the traditional use of the student population, but research has shown there are different characteristics to the student population and research has warned it is not to be described as representative of the general population (Paolacci and Chandler, 2014).

A study was conducted replicating that of Cleveland and McGill (1984) using online methods and crowdsourcing for participants (Heer and Bostock, 2010). The results from Heer and Bostock's (2010) study recreating previous visualisations studies showed that crowdsourced results (via MTurk) provided a good match to previous (in-person) study results. This suggests that by changing the method of implementing questionnaires (i.e., online) does not affect the participants' responses, as the same conclusions were reached when using this methodology.

There were some other conclusions that have been used as a guide in the current research from the Heer and Bostock (2010) research. The study found that increasing chart heights beyond 80 pixels did not increase accuracy (Heer and Bostock, 2010), therefore the charts were sized at 80 pixels for maximum accuracy and to minimise the effects of chart size across different participants and different screens. Heer and Bostock (2010) also found adding gridlines improved accuracy, though post-hoc tests found no significant difference between 10 and 20 gridlines. Concurrent with this, it was concluded that error increased with a chart height of 40 pixels and gridlines of 10 units. To account for the effect of gridlines and the effect gridlines can have on different chart sizes, the present research did not include gridlines. Given there is an effect, this is to allow the research (and participants) to focus on the bars in the bar chart that are highlighted and not use gridlines to answer questions.

3.3 Website

To collect data and run the questionnaire online, a website was designed and published to the Mudfoot Server owned by Manchester Metropolitan University. The website was coded in PHP and HTML which was linked to the SQL database (discussed in Section 3.4). The code used for the website was provided by Dr Kleerekoper who had designed similar websites for studies on visualisation design. Some changes were made to the wording of the questionnaire and changes were made to the demographic questions. WinSCP was used to communicate with the Mudfoot Server and upload PHP/HTML files.

The site opened on an index page detailing the aims of the questionnaire and explaining how the website worked. The start quiz button then leads users to the demographic question, after completion the button used an INSERT statement to submit answers to the SQL database. Each question is then presented with a show visualisation button and answer box. By using a reveal visualisation button (and alerting users to it) the time is as accurate as possible as the question reading time is not included in the time taken to give an answer. To give an answer, a slider and answer box was given, after pressing submit the site moves to the next question or closing page (if ten questions have been completed). The time taken function begins when the show visualisation button is clicked and ends when the submit answer button is clicked.

Following the submit answer button, the site uses a SQL INSERT statement to enter the answers into the database.

3.4 Database

Database was created to store the questions loaded to the website using MySQL Workbench. The database contained 3 tables; vis_question which holds the questions and information about the questions, as well as the answers submitted on the website, vis_participant which holds information about the participant including any feedback provided and demographic information. The entity relationship diagram below shows the design of the database and fields used.

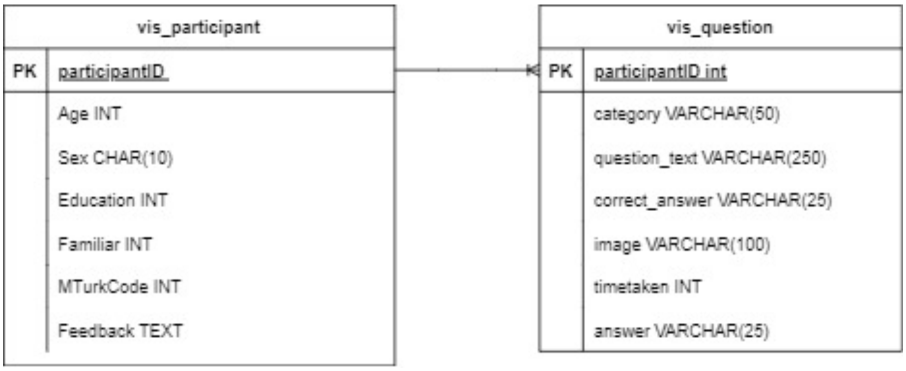


Figure 1: Entity Relationship Diagram

Chapter 4: Experimental Implementation

4.1 Introduction

Chapter 4 will give an overview of the experiment conducted, including detailed descriptions and examples of the three stimulus types and an overview of the procedure following in the experiment.

4.2 Stimuli

In the current study, the chosen emphasis effects are evaluated based on how quickly and accurately a viewer can locate the bar and estimate a value. Two different types of highlighting were used, the first involved changing the colour of the target bar to a brighter colour in comparison to the other bars. The second method involved removing the target bar and replacing it with a line from the y-axis across all parts of the bar chart. The method and justification for the designs are discussed below in Sections 4.2.1, 4.2.2 and 4.2.3.

All stimuli were created using the matplotlib library with Python in the Google Colabatory environment. matplotlib is a Python package for plotting that generates production quality graphs (Ari and Ustazhanov, 2014) and is cited as one of the most popularly used data visualisation libraries of python (Sial, Rashdi and Khan, 2021).

When creating all chart types, a random number array was created with a size of 5 columns and 50 rows (using NumPy package random functions). This meant when creating the charts, the same numbers were used for all charts except the horizontal line method as these were designed differently and 50 charts could be created simultaneously, contributing to experimental design robustness and reproducibility.

4.2.1 Baseline Bar Chart

To begin, 50 arrays of 5 numbers were created at random using the random integer function of NumPy. These arrays were then used in loops to create 50 charts simultaneously. Baseline bar charts were created as a standard measure that the emphasis effects charts could be compared to, therefore all the charts were the same apart from the length of the bars.

The Seaborn set function was used to ensure consistency across all plots. The background was set to white to avoid distraction, and the chart size was set the same for all charts. Gridlines were also removed as the Heer and Bostock (2010) research showed these can have an impact on how charts are read. Each of the scales was fixed to only show between 0 and 100 (in steps of 20) these scales were kept the same across all charts. The seaborn set items used are shown below.

```
sns.set(
    { "figure.figsize": (7,4) },
    color_codes=True
)
sns.set(font_scale=0.8)
sns.set_style("white")
```

4.2.2 Highlighted Bars

Visual Salience as discussed in Chapter 2, refers to how much an item stands out in comparison to neighbouring items. Janicke and Chen (2010) stated that higher levels of visual salience would attract higher attention should be reserved for parts the designer wants to draw attention to, the same has also been said in Mairena et al., (2020) who use the term emphasis effect. The emphasis effect in this instance is the highlighted bar. A bright colour was chosen as this would lead to higher levels of visual salience and utilising colour as this is a preattentively observed attribute (Bartram et al., 2017).

In these charts, the aim was to highlight a single bar and ask the participant to estimate the value of that bar. To ensure the results of the questionnaire were due to the emphasis effects being used, the only thing changed between charts was the random value of the bars.

Each of the charts had 5 bars in total, 1 target and 4 distractors, these bars were named the same in each chart ("A", "B", "C", "D", "E"). No ordering was applied to any of the charts as visual encoding research suggests that participants could have used this to gain other information based on the overall shape of the chart (Kosara, 2022). 50 charts were created with the first column (“A”) as the target bar, 50 were created with the second column (“B”) and 50 were created with the last column (“E”) as the target. Alongside the change in target bar position, the colour of the target bar was also changed. These colours were chosen from a list of matplotlib compatible colours that looked bright in comparison to the grey.

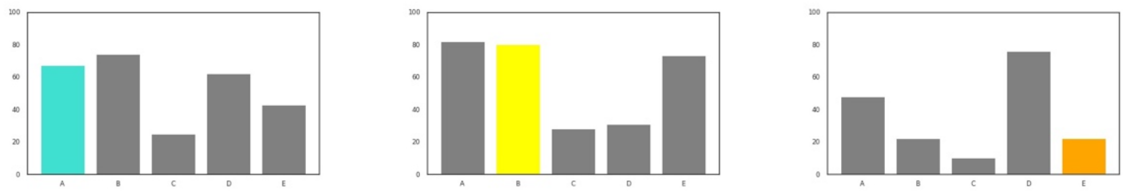


Figure 2: Examples of highlighted column emphasis effect

As described in Chapter 2, changing the colour of a bar is to attract attention to that bar and improve the readability, speed and accuracy of the bar or to imply importance. The aim of this study is to establish if it is more efficient to highlight bars in comparison to a standard bar chart (no highlighting). The above are used as time-invariant highlighted charts and throughout the remaining chapters are referred to as “highlighted bars”.

4.2.3 Horizontal Line Bars

Although there has been research on emphasis effects used in charts covering a variety of preattentive attributes, such as shape, size, blur, hue and orientation, little has been done in other areas. The present research aims to use a horizontal line as an emphasis effect and replace the bar with a line, this aligns with the preattentive processing research as the aim is to attract the attention to the coloured line, however, the bar is removed so this could confuse viewers that are used to looking at the bar for the value.

Visual encoding research suggests that when interpreting a bar chart, a user looks at the length of a bar and estimates a value using the y-axis (Kosara, 2016), the horizontal line chart uses a different type of encoding and allows users to look at the line where it crosses the y-axis, though there is little empirical evidence to suggest this is the case, based on chart research this is what is expected. This type of chart is considered a bar chart, but the target column is now a horizontal line.

Heer and Bostock (2010) found gridlines affect the way participants interpreted charts in the study, gridlines are background lines that stretch from the x-axis tick point to the edge of the chart, making it easier to line up bars with the value of its length on the x-axis. Based on this

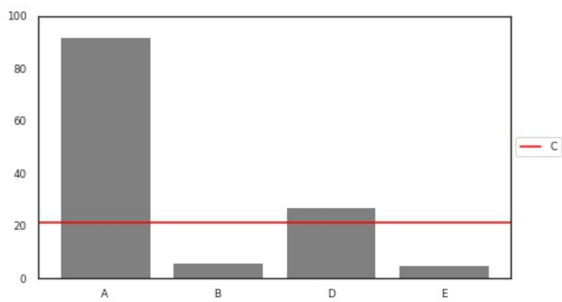


Figure 3: Example of horizontal line emphasis effect

assertion, it is posited in the current research that the horizontal line will allow accurate estimations as it shows the exact value on the x-axis where (ordinarily) a bar would reach. Image NUMBER below is an example of the horizontal line chart.

From the above, the main aims of the chart can be seen. The target bar in this instance is “C”, which has been removed from the x-axis. The alternative is a red horizontal line that represents the top of where the bar would reach. The legend box to the right also denotes the red line representing “C” the target bar. The line has been coloured red to further leverage preattentive processing as this is cited as the most common preattentive attribute to attract attention (Bartram et al., 2017).

The horizontal line charts were designed in the same manner as the standard bar and highlighted bar charts, however, due to the column being removed and replaced with a line the random number size was 50 arrays each containing 4 numbers. A separate random number was then generated and allocated to the horizontal line, representing column “C”.

4.3 Procedure

Participants were recruited via Amazon’s Mechanical Turk to allow for a broader range of participants. When the questionnaire was published, MTurk users access it through the website and complete the 10 questions. At the end of the 10 questions, they are provided with a code which is entered into MTurk to confirm completion and authorise payment. Participants in this study were paid $1.50. The questionnaire took 15 minutes (maximum) to complete, totalling an hourly rate of $6. Research by Heer and Bostock (2010) found that increasing the amount per questionnaire led to a higher number of responses but not necessarily higher quality responses.

When completing the questionnaire, users are first shown an instruction screen (Appendix 1) that describes how the website works, once this screen has been read, the user clicks ‘start quiz’ which leads to the demographics questionnaire (Appendix 2). Participants are told the demographics are required to participate. After the demographic questionnaire, participants were presented with the first question. Participants are instructed to read the question before revealing the visualisation where the answers are taken from, this was to account for reading speed when measuring timed responses, the time measure begins when the button is pressed, after the question is read for all participants. Participants are lead through 10 random questions selected from the database. After 10 are completed an MTurk code is presented and the opportunity to provide feedback via a textbox.

4.4 Statistical Analysis

The aim of the research is to investigate the difference in speed and accuracy of chart interpretation when emphasis effects are used, compared to a standard bar chart. Each of the hypotheses predicts a difference in accuracy and speed for the chart types. A t-test is used to compare the means of two groups and see if the two groups are significantly different. In this study, the null hypothesis is that there is no difference between the mean speed and mean accuracy of participants responses to a bar chart and a chart with an emphasis effect. To reject or not reject the null hypothesis a t-test will be used. The absolute difference will be calculated for each of the participants answers and the correct answer to the question. The t-test will be a two-tailed independent test that will compare the mean difference in the standard bar chart to the mean difference in the emphasis effect charts. The time taken to complete the questions will also be tested. A two-tailed independent t-test will measure the mean difference in response time in each of the emphasis effect charts compared to the baseline chart. The p-value will be used to reject or not reject the null hypothesis, the significance level used for all hypotheses is 5% ($p<0.05$). If $p<0.05$ the null hypothesis can be rejected, if it is greater ($p>0.05$) the null hypothesis cannot be rejected.

If data is not normally distributed, a Mann-Whitney U Test will be used to compare the differences between the two chart types. The Mann-Whitney test does not require normally distributed data. The test used to test the hypotheses will depend on the distribution of the data, this will be determined with a histogram chart. Once the distribution of the data is known, a statistical test will be chosen.

4.4.1 Statistical Power

Statistical power is the probability of detecting an effect, if there is a true effect present to detect or the probability that a test correctly rejects the null hypothesis (Ellis, 2010). The higher the statistical power in an experiment, the lower the chance of a Type II error (false negative). Statistical power can be used to estimate how many participants is suitable for a study, based on effect size, significance level and statistical power. To calculate statistical power for the current study, the default or standard values were used, these are: effect size = 0.8, significance level = 0.05 and power = 0.8 (80%). The statsmodel package was used with Python to calculate the sample size that meets this criterion.

```python
from statsmodels.stats.power import TTestIndPower
# parameters for power analysis
effect = 0.8
alpha = 0.05
power = 0.85
# perform power analysis
analysis = TTestIndPower()
result = analysis.solve_power(effect, power=power, nobs1=None,
ratio=1.0, alpha=alpha)
print('Sample Size: %.3f' % result)
```

The test shows that for a statistical power of 80%, 25 participants are required. The current study recruited 31 participants, suggesting a statistical power closer to 85%.

The below chart shows how the change in effect size and participants could impact the statistical power of the test.

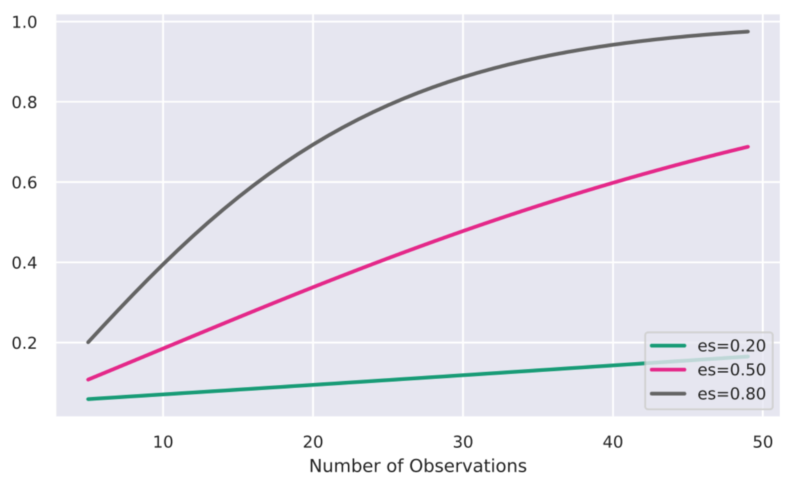


Figure 4: Statistical Power and Effect Size

The chart shows that for an effect size of 0.8 (used in the current research calculation) that the power of the test would begin to plateau at around 50 participants. It also shows lowering the effect size would drastically reduce the power of the test.

Chapter 5: Experimental Results

5.1 Introduction

Chapter 5 will give an overview of the participants in the study including the demographic information. The hypotheses for the study will then be presented with a detailed description of the analyses undertaken to test the hypotheses. The results will then be interpreted and there will be a discussion on if the hypotheses have been proved or not.

5.2 Participant Information

Participants were recruited solely online through Mechanical Turk, there were no entrance criteria and no prerequisites to participate. In total, 31 participants were recruited, the following gives detailed information about the participants.

The ages of the participants were recorded before the questionnaire began. Each participant entered their age in a box. Figure 5 shows a histogram plot of the ages recorded for 31 participants. The youngest participant recorded was 24 years old, and the oldest participant was 48 years old. The histogram plot also shows most participants to be between the ages of 35 and 48 years old. The mean age for all participants was 38.6 (*SD=5.77*) years. The

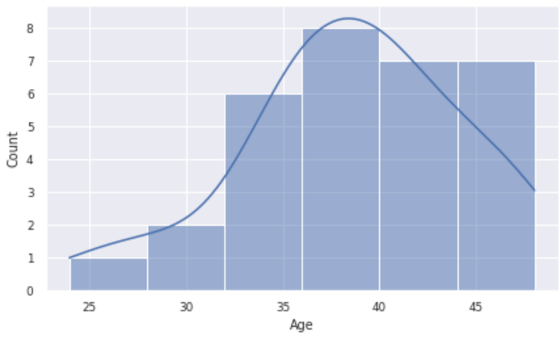


Figure 5: Age distribution

histogram alongside the mean and standard deviation suggests a relatively wide range of participant ages. The analysis conducted on participants in the study can be found at the link in Appendix C.

Participants' gender was also recorded in the demographic information, the options were male, female or non-binary. The table 1 shows the number of participants per gender with the average age of each gender and standard deviation.

	Number of Participants	*Mean Age*	*Standard Deviation*
Male	20	38.05	6.66
Female	10	39.90	5.48
Non-Binary	1	36	0

Table 1: Participant information

Familiarity with visualisations was measured on a three-point Likert scale, a subjective measure decided by the participant, the options were 'a little familiar', 'not very familiar' or 'very familiar'. Figure 6 shows the spread of ages across each familiarity option as well as the gender of the participants. Most participants (19) selected a little familiar, the fewest participants selected not very familiar (4), and the remaining participants selected very familiar (8). The widest spread of ages was in female participants who selected very familiar.

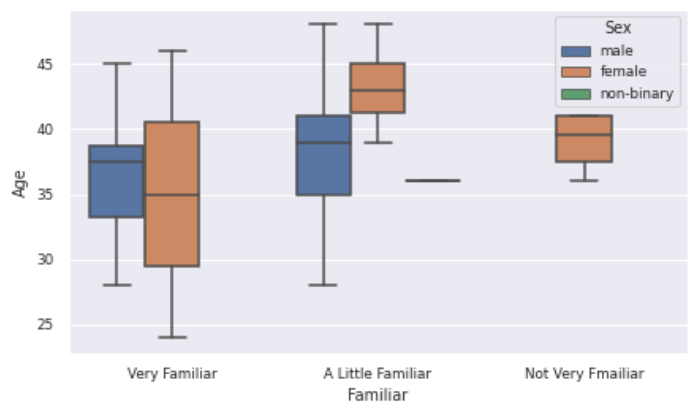


Figure 6: Age and familiarity box plot

Education was also measured in the demographic questionnaire; participants were asked to select their highest qualification from a pre-set list. Given the ubiquitous nature of charts in academic settings (amongst others), it would be expected that those with higher qualifications/longer in education would have more familiarity with charts. Figure 7 shows the education level of participants with a count for which level of familiarity they selected. The chart shows that most participants have a bachelor's degree and most of those with a bachelor's degree selected 'a little familiar' on the self-reported scale. Those who selected the highest familiarity option have a varying level of education, ranging from high school to master's degree. The merits and limitations of using self-reported measures (for familiarity) will be discussed in Chapter 6.

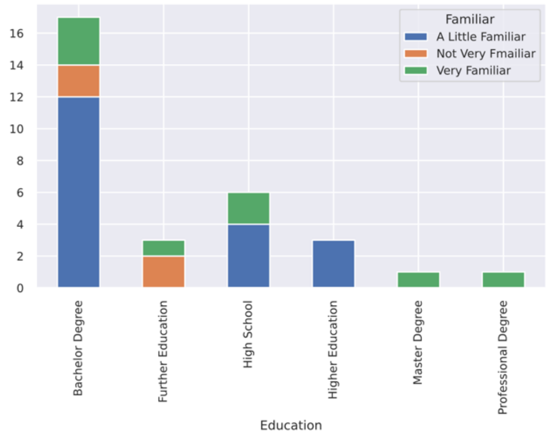


Figure 7: Education and familiarity stacked bar chart

5.3 Data Analysis

5.3.1 Establishing Baseline Responses

Each of the questions was allocated a category, as described in Chapter 3. Firstly, the baseline category was analysed. In total, there was 50 baseline questions answered, of these 9 (18%) were answered correctly. Figure 8 shows the correct answers and participants answer on a scatter plot, the orange points show which were answered correctly. The answer to each question could have been any whole number between 5 and 100, a correct answer is only registered when the answer is guessed exactly.

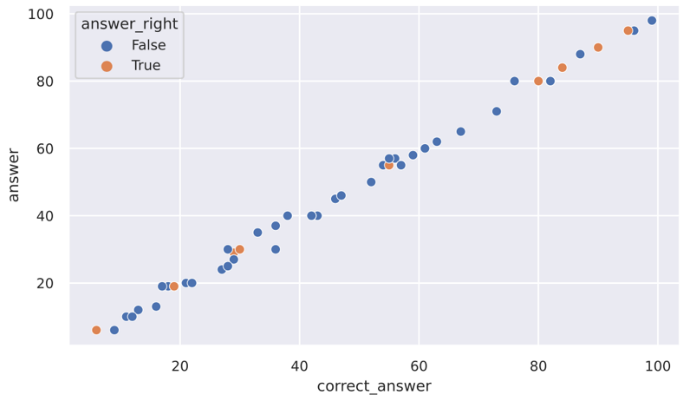


Figure 8: Answer given and correct answer scatter plot

The baseline results were also analysed in relation to the familiarity score given by participants. Given the literature presented in Chapter 2, it would be expected that those with higher graphical literacy would answers more questions correctly. Figure 9 shows those with most familiarity with graphs did not have the most correct answers. The most familiar had 2 correct answers, the least familiar had 1 correct answer and the mid-group had 6 correct answers.

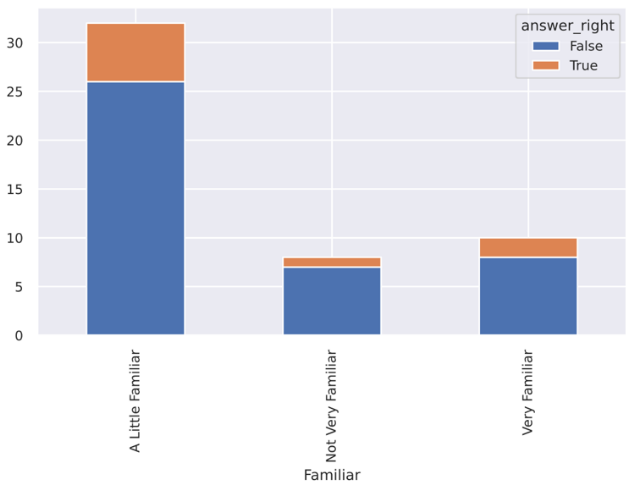


Figure 9: Answer proportion and familiarity

Given the level of accuracy required to get a correct answer, further analysis was conducted to see how many correct answers were given within 3 of the actual correct answer. Figure 10 shows the difference in the number of correct answers when an allowance is given (3 over or under). The chart shows the increase in correct answers, the lowest familiarity group increased from 1 to 8, the highest familiarity increased from 2 to 8 and the mid-group increased from 6 to 29. Based on this increase of correct answers, the difference was calculated to use as an accuracy measure in the hypothesis testing.

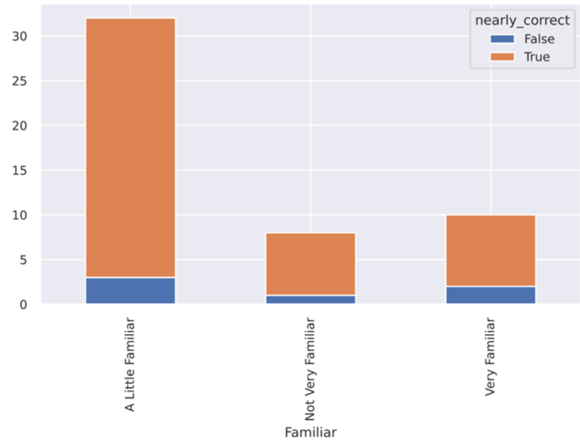


Figure 10: Adjusted correct answer proportion and familiarity

5.3.2 Hypothesis 1

H_0: There will be no rank sum difference between baseline charts and emphasis effect charts.

H_1: Using a highlighted bar as an emphasis effect will result in faster response times from participants, compared to a standard bar chart.

Firstly, a histogram plot was created for the baseline data to see the distribution of the time taken to complete a baseline question. Figure 11 shows the distribution is not normal, therefore a Mann-Whitney U test will be utilised.

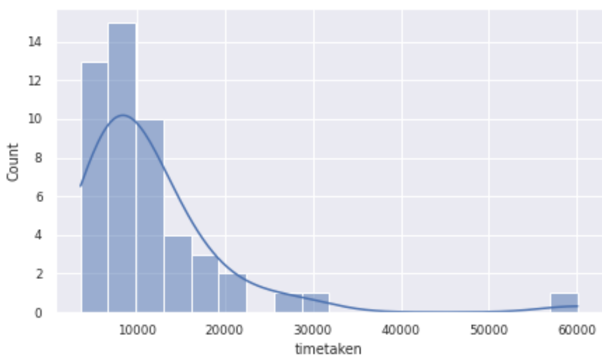


Figure 11: Time Taken for Baseline Questions

Secondly, a histogram plot was created to assess the distribution of the time taken to answer the highlighted bar questions. Figure 12 shows the distribution is not normal.

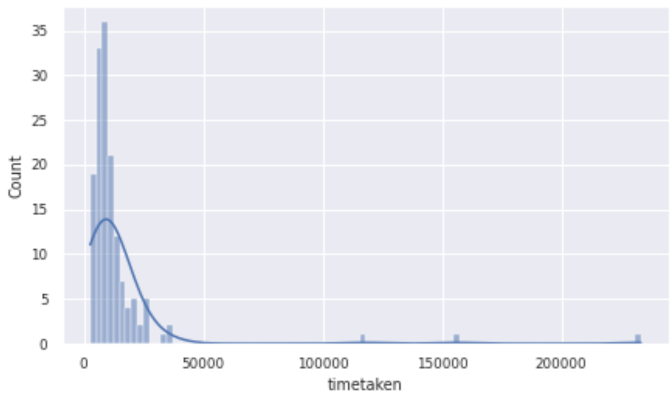


Figure 12: Time Taken for Highlighted Question

To test H_1 a Mann-Whitney test was conducted. The Mann-Whitney test indicated that the median time taken was less for the baseline charts (Mdn=8931.0) than for the highlighted bar charts (Mdn= 9010.5), $U = 3813.5$, $p=.860$. Although the median time was faster for the baseline chart, it was not significantly faster. The null hypothesis cannot be rejected.

5.3.3 Hypothesis 2

H_2: Using a highlighted bar as an emphasis effect will result in more accurate responses from participants, compared to a standard bar chart.

Firstly, as with H_1 histograms were created to assess the distribution of the difference data. Figure 13 shows the lowest difference was 0 and the highest difference, 6.

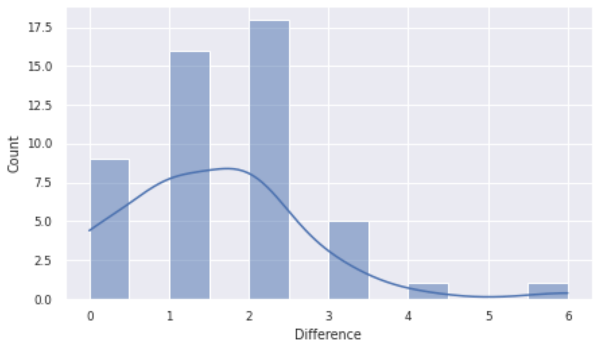
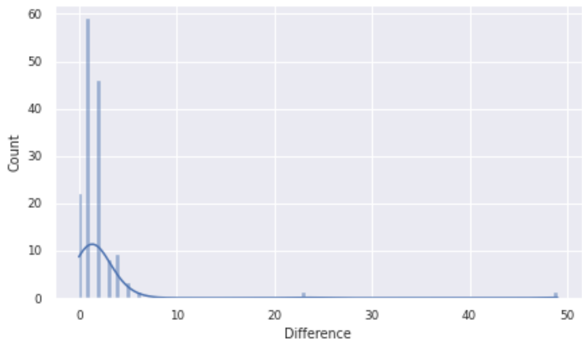


Figure 13: Baseline Difference Distribution

The two bar charts show skewed distribution for both sets of difference data. Some outliers are present in the data, given outliers occurred in all datasets they were left in as there would

Figure 14: Highlighted Bar Difference Distribution



have been less statistical power had participants been removed and it is plausible that charts were read incorrectly, figure 14 illustrates the outliers at around a difference of 50.

To test H_2 a Mann-Whitney test was conducted. The Mann-Whitney test indicated that the median difference was greater for the baseline charts (Mdn=1.5) than for the highlighted bar charts (Mdn= 1.0), *U = 3737.0*, *p=.097*. The results show a large *U* value and a not significant p value. Although the median difference is greater for the baseline chart, it is not significant. The null hypothesis cannot be rejected.

5.3.4 Hypothesis 3

H_3: Using a horizontal line to replace a bar will lead to more accurate responses from participants, compared to a standard bar chart.

As with hypothesis 1 and 2, a histogram plot was created to show the distribution of the difference data of the horizontal line chart.

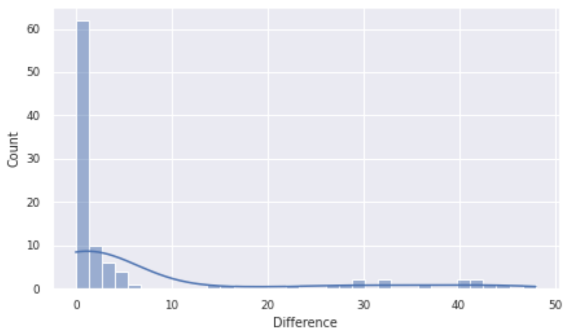


Figure 16: Distribution of Horizontal Line Differences

Figure 15 shows the data is skewed and there is a wide spread of the difference values, ranging from 0 to 50.

To test H_3 a Mann-Whitney test was conducted. The Mann-Whitney test indicated that the median difference was greater for the baseline charts (Mdn=1.5) than for the horizontal bar charts (Mdn= 1.0), $U = 2538.0$, $p=.882$. The results show there is no significant difference in the interpretation of horizontal line charts compared to the baseline charts. The null hypothesis cannot be rejected.

Following on from H_3, an additional test was undertaken to see if the response time was faster on the standard bar chart or horizontal bar chart. A Mann-Whitney test was conducted. The Mann-Whitney test indicated that the median difference was greater for the baseline charts (Mdn=8931.0) than for the horizontal bar charts (Mdn= 9663.0), $U = 2121.5$, $p=.133$.

The Mann-Whitney *'U'* values shows the difference between the two rank totals. The smaller the U value the less likely the event occurred by chance, given the large U value it is likely the effect happened by chance and the difference between the groups is not significant. The null hypothesis cannot be rejected.

5.3.5 Familiarity with Charts

Given the research presented on graphical literacy and its importance to visualisation research, it was hypothesised that higher levels of chart familiarity would lead to more correct responses. Table 2 shows the number of participants who self-reported for each level of familiarity and the percentage of correct answers for each familiarity level. The percentage of correct answers increases as the familiarity level increase, suggesting a relationship. However, there are considerably more participants describing themselves as a little familiar than the other levels.

	Not Very Familiar	*A Little Familiar*	*Very Familiar*
Number of Participants	4	19	8
Percentage of Correct Answers	17.5%	18.4%	20%

Table 2: Familiarity and Correct Answers

5.4 Analysis Summary

The above presents the results of 3 hypotheses based on the research outlined in Chapter 2. It was expected there would be a significant difference in speed of answers and accuracy of answers between a standard bar chart and a chart with emphasis effects due to the preattentive attributes being faster to process. Each of the hypotheses will now be discussed in turn.

Hypothesis 1 stated using a highlighted bar as an emphasis effect will result in faster response times from participants, compared to a standard bar chart. Statistical testing found the median response times for the baseline chart were slightly lower, therefore faster, than the highlighted bar chart. This could be due to the number of outliers in the data which were not removed. These were not removed as the statistical power would have been reduced, however given there is no significant results a reduction in statistical power may have been acceptable.

Hypothesis 2 stated using a highlighted bar as an emphasis effect will result in more accurate responses from participants, compared to a standard bar chart. Statistical testing showed a lower median difference for the highlighted bar than the standard bar, however there was not enough difference to reject the null hypothesis. The lower median difference in the highlighted bar suggests there could be an effect of using emphasis effects which could be investigated further. Further research into this phenomenon should consider the removal of outliers in the data and an increase in participants to accommodate the reduction of data in statistical power testing.

Hypothesis 3 stated using a horizontal line to replace a bar will lead to more accurate responses from participants, compared to a standard bar chart. Although the median difference was lower for the horizontal line chart, the difference was not significantly different, therefore the null hypothesis could not be rejected. Although the test showed no significance there was a lower median difference suggesting further investigation is warranted.

Chapter 6: Evaluation

6.1 Introduction

The following chapter will evaluate the present research and discuss if the aims and objectives outlined in Section 1.2.1 were achieved and to what extent. Firstly, the aims and objectives will be evaluated, in turn, in relation to the research that has been presented in Chapter 1-5. This will be followed by a critical evaluation of each chapter presented.

6.2 Evaluation of Present Research

The first objective was to contribute to literature on data visualisation design. Overall, the project has achieved its aims, as a literature review has been presented with hypotheses based on previous works. This was then followed by an experiment to test the hypotheses, which could inform future works and design. Although the literature review covered various relevant areas of visual perception, the main theories of visual perception were not discussed. There are many theories arising from cognitive psychology (Eysenck and Keane, 2015) that discuss visual perception and how different elements are perceived. The present literature review focussed only on the research and elements related to data visualisation design and interpretation. Further detail on the theories could lead to more ideas being generated to link the fields of visual perception and visualisation design. These can be found in detail in other published works but their inclusion in this paper could have explained the topic in further detail to facilitate a more in depth understanding of how the fields are linked.

Secondly, the research aimed to investigate the use of preattentive processing in visualisation design. This was achieved in Chapter 2 where visual perception, preattentive processing and visual encoding were discussed in detail and was linked to visualisation design research. The current research aimed to draw on previous visual perception research and visualisation research to link the two together based on senior researcher's advice (Kosara, 2022). In part, this was achieved due to the level of discussion around visual perception and how this links to visualisation design. Furthermore, the research was based on the link between these two fields and the experiment was designed accordingly. However, although these fields were discussed and linked together, the only link in the experiment was the design of the charts which were based on preattentive attributes, no research was conducted on attention. There was no empirical measurement of attention in the present study, the stimuli were designed based on previous research, but there is no way of knowing if the effects were registered preattentively or not. This also leads to the discussion of visual encoding, which as discussed in Chapter 2 assumes that charts are interpreted based on encoded elements, but it has been argued that there are other encoded elements at work, but differentiating the two can be difficult (Kosara, 2022). The current research could not account for this.

Thirdly, the research aimed to assess the effectiveness of emphasis effects in bar charts and conduct an experiment to evaluate the effectiveness of emphasis effects in bar charts. This was achieved through the experiment conducted and presented in Chapter 5. Each of the hypotheses were statistically tested and designed to assess effectiveness as an overall measure of speed and accuracy. The speed was tested with Mann-Whitney tests, comparing the baseline chart to the emphasis effect charts, each hypothesis was evaluated, and justification given as to whether they could be rejected. Although none of the tests showed significant differences, some differences were recorded between the groups suggesting there could be an effect. Though each hypothesis was designed based on research, there are many other areas that have not been discussed in the present research. The present research focussed solely on bar charts because they are the most common. However, other research has investigated the emphasis effects on scatter plots (Mairena et al., 2020). To improve the current study, different chart types could have been used, and compared to one another rather than comparing only bar charts, this would lead to more design suggestions for practitioners. Further to this, there are many more emphasis effects than those tested in this study, such as, time-variant effects (motion) and effects that alter the whole chart rather than one element (blurring), including more charts could have led to more design recommendations. However, the second effect presented in this study (horizontal line) is a new concept that has little research, this study hopes to be a preliminary study on this effect and its usefulness.

Another important area to visualisation design discussed in depth in Chapter 2 is that of graphical literacy. This topic is commonly associated with visualisation design and has been discussed in depth in many published works. The present study aimed to use this information to draw further conclusions from the data about the impact of graphical literacy on accuracy and speed of interpretation. However, when designing the study, a Likert scale was used. Likert scales are commonly used to measure attitudes in social science research and other domains; however, they have been widely criticised by others (Willits, Theodori and Luloff, 2016). There is much debate about the number of items that should be included in a Likert scale, in this research the aim was to get an overall view of a participant's graphical familiarity so only three options were given, these were not aggregated or summed as only one question was used. Further to the design around Likert scales is the debate of using self-report scales as measures. It is thought they can be unreliable, and participants can downplay or overplay their actual skill level, therefore the accuracy of this familiarity score cannot be determined.

Finally, the research aimed to make suggestions to visualisation design based on this preliminary research on future emphasis effects. The present study has not made any recommendations on visualisation design for various reasons. Firstly, graphical literacy was not tested in the hypotheses, so it cannot be known if these emphasis effects are effective across the whole population. Similarly, race/ethnicity/location was not recorded so it is not possible to know if different parts of the population interpret different emphasis effects differently. For example, where colour was used there are cultural and societal expectations of colour, such as, red can mean anger in some cultures and happiness/luck in others. Therefore, the colour being used could affect how participants interpreted the value.

Overall, the current research has achieved most of its aims in a robust manner and can be used as preliminary basis for further research. There were some limitations in the study design and further topics could have been included in the literature review. Regardless of the limitations discussed, hypotheses were designed and tested and will contribute to the empirical evidence base of visualisation design.

Chapter 7: Conclusion

6.1 Introduction

This research project has discussed visual perception and visualisation design alongside the importance of linking these fields to better inform design choices. The study has tested multiple hypotheses based on a baseline design and charts with emphasis effects. The project has achieved most of its aims outlined in Chapter 1 and evaluated in Chapter 5. The following are the conclusions from the research, the limitations of the current research and suggestions for future research based on what has been found in the current research.

6.2 Limitations

Although the research carried out is has achieved some important aims, there are some limitations beyond those discussed in Chapter 5 that will now be addressed. Firstly, when designing the study, it became apparent from other research that measuring the magnitude of difference between colours is not possible (Mairena et al., 2020) making visual saliency calculations unattainable. By looking at a chart it is possible to see the difference due to a change in colour, however, when creating empirical evidence an empirical measure should be used to further quantify the changes. For example, if it was possible to know changing the colour within a certain parameter benefitted the user then plateaued, this could be implemented into visualisation design. Similarly, it is not possible to say that brighter colours are better because what is deemed brighter was not quantified, colours were chosen based on visual appeal and differences only. Further to this, as discussed in Chapter 5, when choosing colours it is important to consider the social and cultural implications each colour could have and how this would impact the interpretation of the numbers.

Crowdsourcing participants has proven to be an efficient way to recruit participant at a low cost and high volume. Heer and Bostock (2010) showed the replication of an in-person study via crowdsourcing methods produced the same results, however it can also lead to an increase in individual differences that cannot be accounted for, such as monitor and browser settings. However, this variance also covers a wider variety of realistic viewing scenarios allowing for more general predictions about how viewers will perceive visualisations in the real world - often with higher accuracy than models and results from traditional studies (Mairena et al., 2022). A large criticism of visualisation studies is in the participant recruitment due to an overreliance on the student population (Lam et al., 2013). Lam et al., (2013) also posit that one way to increase validity in visualisation studies is to increase the realism of the task, a reader would likely be interpreting charts on their own browser, therefore the realism is increased. Increasing the realism of the study could have been achieved by using real-world data (Lam et al., 2013), however, there is no way to reliably account for the learning effect, for example some participants may have more experience with the real-world dataset than others which would not be captured in the study other than by an increase in accuracy or speed of interpretation. To increase realism the present study could have used real-world data alongside the random number data and compared the two to see if there were any differences in interpretation.

The present research focused solely on bar charts and emphasis effects on bar charts. To better test emphasis effects, the same approaches (or similar) could be applied to different chart types to see if the emphasis effect is consistent across multiple chart types, something that has not been attempted but is recommended (Mairena et al., 2022). As well as different chart types, there are also multiple emphasis effects that have not been substantially researched, by adding more emphasis effects to the present research, the link between preattentive processing and effective visualisations could have been explored further. However, the present study did use two different types for comparison and have highlighted the importance of preattentive attributes when designing visualisations.

Finally, the familiarity question used in the questionnaire is not a reliable measure of graphical literacy, to build on the results presented the present study could have included a more robust way of measuring this and compared results across varying levels of graphical literacy which would have contributed further to the research area. Jena et al., (2021) discuss the social and ethical implications of excluding groups from visualisation research. The current study could have contributed to this argument by collecting more detailed demographic information and using this to draw comparisons across different populations to improve design for different audiences, particularly those with more recent access to data visualisations via the internet.

6.3 Further Research Recommendations

Building from this research, there are some recommendations for future research. Firstly, as discussed in Section 6.2, measuring demographic information could help inform how different readers interpret the information in visualisations. Further informing chart design with this information could improve readability for previously marginalised groups, such as those who are not taught data skills (Kennedy and Hill, 2018).

Future research should also consider graphical literacy as an important part of the outcomes. Previous research has shown the differing abilities can lead to marginalisation (Kennedy and Hill, 2018), various disadvantages (Zarcostas, 2020) and uncertainty (Kennedy et al., 2020), closing this gap in research can improve outcomes for all readers of visualisations regardless of background or graphical literacy levels. To measure this accurately, research should include a measure of graphical literacy based on a score from various questions of understanding charts, such as, mathematical skills and reading abilities. Correlations could then be accurately undertaken by examining a graphical literacy score in relation to chart accuracy.

Finally, research should continue to examine different types of emphasis effects on different chart types, for example, highlighting used in a bar chart could have a greater effect than if used on a pie chart or could change the way a pie chart is read, whereas there is no evidence of a change in interpretation with highlighted bar charts. Eye-tracking methods are often used in visualisations, this would be a good way to see if the eye is drawn to the preattentive attributes and could further solidify this research. This could also link in with graphical literacy, as it is thought that one of the stages is visual encoding, using this information to see how different emphasis effects are encoded to lead to developments with chart design.

6.4 Summary

This research aimed to assess the impact of using emphasis effects when designing bar charts. Based on quantitative analysis of the present study, there is not enough evidence to state there is a difference between the two chart types. The three research questions presented in Chapter 1 will now be addressed in turn and answered based on the content of this project.

Do readers extract information from a bar faster and more accurately when emphasis effects are used?
To answer this question, statistical testing was used to identify the median difference between a standard bar chart and two different chart types with emphasis effects utilised. There was no significant difference between a standard bar and a highlighted bar chart though the median difference was lower on the highlighted bar chart. This research question was addressed as previously published works suggest using emphasis effects in charts leads to more accurate responses (Mairena et al., 2022). this study cannot empirically support this assertion, however there are some parts of this research that could be useful when conducting further research, such as the crowdsourcing methodology and the use of a baseline measure to test emphasis effects.

Is there a difference in speed and accuracy between a standard bar and one with emphasis effects?
Given the data collected was not normally distributed, statistical tests were used that examined the rank sum median difference of the groups evaluated. The results showed there was a slight difference in speed, the baseline chart was interpreted faster than the highlighted bar chart and horizontal line bar chart. This could be expected for the horizontal line chart as this is a new way of reading bar chart that goes against the widely accepted visual encoding method for bar charts, in that the value is represented by a vertical bar. However, it was expected that changing the colour of a column would make a reader look preattentively and

therefore answer faster as the search time for the target bar is faster. Each highlighted chart featured a highlighted bar either for the first column, third column or final column, these were not investigated separately to assess if the position of target bar influences the speed or accuracy of interpretation.

The literature presented in this study suggested participants would have faster and more accurate responses when interpreting charts with emphasis effects employed. Hypotheses were created in line with research on preattentive processing and data visualisation design, however no significant effects were found in the present study. Although there were no significant differences, it is premature to state there are no differences in charts when emphasis effects are utilised, there was some difference shown in statistical testing, but not at a significant level. As well as the wealth of research that has previously found differences in charts using emphasis effects. Further research may be able to identify these effects more clearly with a different study design or testing.

References

Airaldi, A. L., et al. (2022). "Narrative Visualizations Best Practices and Evaluation: A Systematic Mapping Study."

Barrera-Leon, L., et al. (2020). Systematic variation of preattentive attributes to highlight relevant data in information visualization. 2020 24th International Conference Information Visualisation (IV), IEEE.

Bartram, L., et al. (2017). Affective Colour in Visualization. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. Denver, Colorado, USA, Association for Computing Machinery**:** 1364–1374.

Bateman, S., et al. (2010). Useful junk? The effects of visual embellishment on comprehension and memorability of charts. Proceedings of the SIGCHI conference on human factors in computing systems.

Boy, J., et al. (2014). "A principled way of assessing visualization literacy." IEEE transactions on visualization and computer graphics **20**(12): 1963-1972.

Boyd, D. and K. Crawford (2012). "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." Information, Communication & Society **15**(5): 662-679.

Buhrmester, M., et al. (2016). "Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data?".

Burns, A., et al. (2020). How to evaluate data visualizations across different levels of understanding. 2020 IEEE Workshop on Evaluation and Beyond-Methodological Approaches to Visualization (BELIV), IEEE.

Chen, C. (2010). "Information visualization." Wiley Interdisciplinary Reviews: Computational Statistics **2**(4): 387-403.

Chen, H. M. (2017). "Information visualization principles, techniques, and software." Library technology reports **53**(3): 8-16.

Chen, M. and L. Floridi (2013). "An analysis of information visualisation." Synthese **190**(16): 3421-3438.

Chen, M., et al. (2014). What is visualization really for? The Philosophy of Information Quality, Springer**:** 75-93.

Cleveland, W. S. and R. McGill (1984). "Graphical perception: Theory, experimentation, and application to the development of graphical methods." Journal of the American statistical association **79**(387): 531-554.

Correll, M., et al. (2020). Truncating the y-axis: Threat or menace? Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

Dick, M. (2015). "Just fancy that: An analysis of infographic propaganda in The Daily Express, 1956–1959." Journalism Studies **16**(2): 152-174.

Ellis, P.D., 2010. The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results. Cambridge university press.

Elzer, S., et al. (2011). "The automated understanding of simple bar charts." Artificial Intelligence **175**(2): 526-555.

Etchebehere, S. and E. Fedorovskaya (2017). "Psychophysical study of visual saliency of different hues." Journal of Vision **17**(7): 62-62.

Eysenck, M. W. and M. T. Keane (2015). Cognitive psychology: A student's handbook, Psychology press.

Few, S. and P. Edge (2008). "What ordinary people need most from information visualization today." Perceptual Edge: Visual Business Intelligence Newsletter: 1-7.

Firat, E. E., et al. (2022). "Interactive visualization literacy: The state-of-the-art." Information Visualization **21**(3): 285-310.

Franconeri, S. L., et al. (2021). "The science of visual data communication: What works." Psychological Science in the public interest **22**(3): 110-161.

Fronza, I., et al. (2013). Cooperation wordle using pre-attentive processing techniques. 2013 6th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), IEEE.

Glazer, N. (2011). "Challenges with graph interpretation: A review of the literature." Studies in science education **47**(2): 183-210.

Hall, K. W., et al. (2016). "Formalizing Emphasis in Information Visualization." Computer Graphics Forum **35**(3): 717-737.

Harsh, J. A., et al. (2019). "“Seeing” data like an expert: An eye-tracking study using graphical data representations." CBE—Life Sciences Education **18**(3): ar32.

Healey, C. G., et al. (1995). "Visualizing real-time multivariate data using preattentive processing." ACM Transactions on Modelling and Computer Simulation (TOMACS) **5**(3): 190-221.

Heer, J. and M. Bostock (2010). Crowdsourcing graphical perception: using mechanical turk to assess visualization design. Proceedings of the SIGCHI conference on human factors in computing systems.

Jänicke, H. and M. Chen (2010). A salience-based quality metric for visualization. Computer Graphics Forum, Wiley Online Library.

Jena, A., et al. (2021). "The next billion users of visualization." IEEE Computer Graphics and Applications **41**(2): 8-16.

Kennedy, H. and R. L. Hill (2018). "The feeling of numbers: Emotions in everyday engagements with data and their visualisation." Sociology **52**(4): 830-848.

Kennedy, H., et al. (2016). "The work that visualisation conventions do." Information, Communication & Society **19**(6): 715-735.

Kennedy, H., et al. (2020). "Data visualization and transparency in the news." Data visualization in society: 169-185.

Kosara, R. (2016). An empire built on sand: Re-examining what we think we know about visualization. Proceedings of the sixth workshop on beyond time and errors on novel evaluation methods for visualization.

Kosara, R. (2022). "More Than Meets the Eye: A Closer Look at Encodings in Visualization." IEEE Computer Graphics and Applications **42**(2): 110-114.

Kosara, R., et al. (2003). "Thoughts on user studies: Why, how, and when." IEEE Computer Graphics and Applications **23**(4): 20-25.

Krekhov, A., et al. (2019). "Deadeye visualization revisited: Investigation of preattentiveness and applicability in virtual environments." IEEE transactions on visualization and computer graphics **26**(1): 547-557.

Krekhov, A. and J. Krüger (2018). "Deadeye: A novel preattentive visualization technique based on dichoptic presentation." IEEE transactions on visualization and computer graphics **25**(1): 936-945.

Lallé, S., et al. (2016). "Prediction of individual learning curves across information visualizations." User Modelling and User-Adapted Interaction **26**(4): 307-345.

Lam, H., et al. (2011). "Empirical studies in information visualization: Seven scenarios." IEEE transactions on visualization and computer graphics **18**(9): 1520-1536.

Lewandowsky, S. and I. Spence (1989). "The perception of statistical graphs." Sociological Methods & Research **18**(2-3): 200-242.

Liang, J. and M. L. Huang (2010). Highlighting in information visualization: A survey. 2010 14th International Conference Information Visualisation, IEEE.

Mairena, A., et al. (2020). "A baseline study of emphasis effects in information visualization."

Manovich, L. (2011). "What is visualisation?" Visual Studies **26**(1): 36-49.

Mast, F. W. and S. M. Kosslyn (2002). "Eye movements during visual mental imagery." Trends in Cognitive Sciences **6**(7): 271-272.

Mayer-Schönberger, V. and K. Cukier (2013). Big data: A revolution that will transform how we live, work, and think, Houghton Mifflin Harcourt.

Meyer, M. and J. Dykes (2019). "Criteria for rigor in visualization design study." IEEE transactions on visualization and computer graphics **26**(1): 87-97.

Mogull, S. A. and C. T. Stanfield (2015). Current use of visuals in scientific communication. 2015 IEEE international professional communication conference (IPCC), IEEE.

Nærland, T. U. and M. Engebretsen (2021). "Towards a critical understanding of data visualisation in democracy: a deliberative systems approach." Information, Communication & Society: 1-19.

Niño Zambrano, R. and Y. Engelhardt (2008). Diagrams for the masses: Raising public awareness–from Neurath to Gapminder and Google Earth. International conference on Theory and application of diagrams, Springer.

Okan, Y., et al. (2016). "How people with low and high graph literacy process health graphs: Evidence from eye-tracking." Journal of Behavioral Decision Making **29**(2-3): 271-294.

Okan, Y., et al. (2012). "When higher bars are not larger quantities: On individual differences in the use of spatial information in graph comprehension." Spatial Cognition & Computation **12**(2-3): 195-218.

Padilla, L. M., et al. (2019). "Toward objective evaluation of working memory in visualizations: A case study using pupillometry and a dual-task paradigm." IEEE transactions on visualization and computer graphics **26**(1): 332-342.

Paolacci, G. and J. Chandler (2014). "Inside the Turk: Understanding Mechanical Turk as a participant pool." Current directions in psychological science **23**(3): 184-188.

Pietz, J., et al. (2020). "Chasing John Snow: Data analytics in the COVID-19 era." European Journal of Information Systems **29**(4): 388-404.

Robinson, A. C. and G. Center (2006). Highlighting techniques to support geovisualization. Proceedings of the ica workshop on geovisualization and visual analytics, Citeseer.

Rodrigues, J. F., et al. (2008). The Visual Expression Process: Bridging Vision and Data Visualization, Berlin, Heidelberg, Springer Berlin Heidelberg.

Rosli, M. H. W. and A. Cabrera (2015). "Gestalt principles in multimodal data representation." IEEE Computer Graphics and Applications **35**(2): 80-87.

Sadiku, M., et al. (2016). "Data visualization." International Journal of Engineering Research And Advanced Technology (IJERAT) **2**(12): 11-16.

Seong, Y., et al. (2020). Guidelines for cybersecurity visualization design. Proceedings of the 24th Symposium on International Database Engineering & Applications.

Skau, D. and R. Kosara (2016). Arcs, angles, or areas: Individual data encodings in pie and donut charts. Computer Graphics Forum, Wiley Online Library.

Stalph, F. and B. Heravi (2021). "Exploring data visualisations: An analytical framework based on dimensional components of data artefacts in journalism." Digital Journalism: 1-23.

Stofer, K. A. (2016). "When a picture isn't worth 1000 words: Learners struggle to find meaning in data visualizations." Journal of Geoscience Education **64**(3): 231-241.

Tufte, E. R. (1983). The visual display of Quantitative Information.

Waldner, M., et al. (2017). Exploring visual prominence of multi-channel highlighting in visualizations. Proceedings of the 33rd Spring Conference on Computer Graphics. Mikulov, Czech Republic, Association for Computing Machinery**:** Article 8.

Wang Baldonado, M. Q., et al. (2000). Guidelines for using multiple views in information visualization. Proceedings of the working conference on Advanced visual interfaces.

Ware, C. (2019). Information visualization: perception for design, Morgan Kaufmann.

Waskom, M. L. (2021). "Seaborn: statistical data visualization." Journal of Open Source Software **6**(60): 3021.

Willits, F.K., Theodori, G.L. and Luloff, A.E., 2016. Another look at Likert scales. *Journal of Rural Social Sciences*, *31*(3), p.6.

Yalcinkaya, M. and V. Singh (2019). "Exploring the use of Gestalt's principles in improving the visualization, user experience and comprehension of COBie data extension." Engineering, Construction and Architectural Management.

Zarocostas, J. (2020). "How to fight an infodemic." The lancet **395**(10225): 676.

Zhao, M., et al. (2019). Neighbourhood perception in bar charts. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.

Appendix A

The instructional page presented to participants before beginning the quiz.

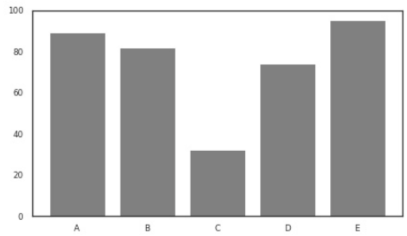
Data Visualization Quiz

Welcome to our Data Visualization Quiz

The aim of this quiz is to gather empirical evidence to test a small number of hypotheses in the field of data visualization. Thank you for taking part. Please note the following important points:

- The quiz contains a set of simple questions which, if answered properly, should take around 10 mins (HIT is limited to 25 mins).
- Each question will display on its own page. At first you will see the question but no visualisation (table, bar chart etc). You will have to press a button to reveal the visualisation and the answer box. Please do not click the "reveal" button until after you have read the question, otherwise our timing results will be affected.
- **DO NOT** press the back button at any point during the quiz as this will reset your progress.
- Please pay close attention to each question until the quiz is finished. **ALL** questions must be answered in a reasonable time, if any question is answered too slowly (based on our pilot studies) you will not be given the MTurk quiz completion code at the end.
- The first question of the quiz will ask for some demographic information (Age, Sex, Education Level, Familiarity with Charts) and by taking the quiz you agree to provide this information. Some of the questions are taken from the USA 2020 Census. We will not capture or store any personally identifiable information.

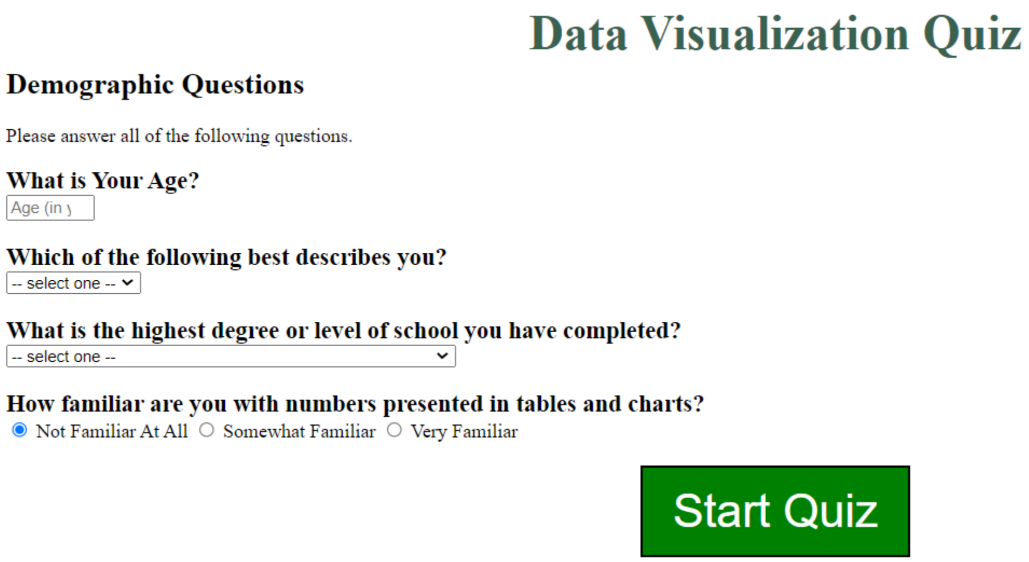
The screenshot below shows how a typical question might look



Start Quiz

Appendix B

Demographic page from the quiz website.



Data Visualization Quiz

Demographic Questions

Please answer all of the following questions.

What is Your Age?

Age (in y

Which of the following best describes you?

-- select one --

What is the highest degree or level of school you have completed?

-- select one --

How familiar are you with numbers presented in tables and charts?

Not Familiar At All Somewhat Familiar Very Familiar

Start Quiz

Appendix C

Google Collaboratory Notebooks.

Notebook 1: Charts created and used in experiments.

https://colab.research.google.com/drive/1V_WbSJP87-08hxOMYkTYv4cBBBHBgaDs?usp=sharing

Notebook 2: Analysis notebook and analysis charts, including rough work.

https://colab.research.google.com/drive/14SkFuZiCw18ida2W80cTAiFlg-lbOYYJ?usp=sharing