TU WIEN Informatics

Quantitative Auswertung von Lesegeschwindigkeit und Fehlerquote bei der Interpretation von visuellen Inhalten

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Visual Computing

eingereicht von

Michaela Tuscher, BSc Bakk.rer.soc.oec.
Matrikelnummer 00827032

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Univ.Prof. Dipl.-Ing. Dr.techn. Eduard Gröller
Mitwirkung: Univ.Lektorin Dipl.-Ing. Dr.techn. Johanna Schmidt

Wien, 6. März 2023

Michaela Tuscher	Eduard Gröller

Quantitative Evaluation of Reading Times and Error Rates When Interpreting Visual Content

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Visual Computing

by

Michaela Tuscher, BSc Bakk.rer.soc.oec.
Registration Number 00827032

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Dipl.-Ing. Dr.techn. Eduard Gröller
Assistance: Univ.Lektorin Dipl.-Ing. Dr.techn. Johanna Schmidt

Vienna, 6th March, 2023

Michaela Tuscher | Eduard Gröller

Erklärung zur Verfassung der Arbeit

Michaela Tuscher, BSc Bakk.rer.soc.oec.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 6. März 2023

Michaela Tuscher

Acknowledgements

First of all, I want to thank my boyfriend, parents, and cats for their support.

I also want to thank Eduard Gröller and, especially, Johanna Schmidt for their support in writing this thesis and providing me with valuable feedback.

Further, I want to extend a big thank you to all of the people that dedicated their time to participate in the pre-study and the exploratory study.

This work was enabled by the Competence Centre VRVis. VRVis is funded by BMK, BMDW, Styria, SFG, Tyrol, and Vienna Business Agency in the scope of COMET - Competence Centers for Excellent Technologies (879730) which is managed by FFG.

The research was funded by the Austrian FFG FemTech program (project no. 350999).

Kurzfassung

Statische Visualisierungen wie beispielsweise Diagramme, Landkarten oder Zeichnungen spielen heutzutage eine immer größere Rolle im täglichen Leben. Daher ist es wichtig herauszufinden, wie hilfreich sie wirklich sind - besonders im Vergleich zu Texten. Diese Diplomarbeit beschäftigt sich mit der Frage, ob Visualisierungen oder Texte schneller von Menschen erfasst werden können, und welche Darstellungsform im Besonderen besser verständlich ist. Zu diesem Zweck führten wir eine explorative Studie durch, in der Verarbeitungszeit und Fehlerrate bei der Interpretation von Texten und Visualisierungen gemessen wurden.

Durch eine Vorstudie, bei der die Testpersonen jeweils vier Visualisierungen in ihren eigenen Worten beschreiben mussten, fanden wir heraus, worauf Leute in Visualisierungen am meisten achten. In den beschreibenden Texten der Testpersonen lag der Fokus einerseits auf Extremwerten der Visualisierungen, andererseits aber auch auf interessanten anderen Werten, die keine Extremwerte darstellten. Darüber hinaus wurden in den Texten oft verschiedene Werte verglichen um die Visualisierung zu beschreiben. Durch diese Erkenntnisse aus der ersten Studie konnten wir die Texte für unsere darauffolgende vergleichende Studie bestmöglich verfassen.

Für die eigentliche explorative Studie zum Erfassen von Verarbeitungszeit und Fehlerrate im Interpretieren von visuellen Inhalten verwendeten wir 15 Visualisierungen und Texte, die jeweils die gleichen Informationen enthielten. Jede Testperson erhielt mindestens sechs Aufgaben, davon mindestens drei Visualisierungen und drei Texte, zu denen die Testperson jeweils drei Fragen beantworten musste. Während der Bearbeitung der Beispiele wurde die Zeit gemessen. Ein interessantes Ergebnis war, dass die Testpersonen beim Bearbeiten von Aufgaben mit Visualisierungen durchschnittlich um etwa 1,3 mal schneller waren als mit Texten. Dabei handelt es sich um einen statistisch signifikanten Unterschied. Allerdings konnten wir keine signifikanten Unterschiede bei der Fehlerrate zwischen Aufgaben mit Texten und Aufgaben mit Visualisierungen feststellen. Außerdem stellten wir bei den Beispielen mit Texten Korrelationen zwischen Textlänge und Bearbeitungsgeschwindigkeit sowie Textlänge und Fehlerrate fest. Wir konnten keinen Einfluss der Inhalte der Visualisierungen und Texte auf die Geschwindigkeit oder Fehlerrate feststellen. Allerdings fanden wir Fälle, in denen Beispiele mit Visualisierungen um mehr als 50 % schneller bearbeitet wurden als die dazugehörigen Texte.

Unsere Ergebnisse bieten nun eine solide Basis zur Definition weiterer Hypothesen hinsichtlich der Lesbarkeit von Visualisierungen im Vergleich zu Text. In dieser Arbeit präsentieren wir die finalen Hypothesen, welche aus unserer explorativen Studie hervorgehen. Wir betrachten diese als äußerst interessant für die Visualisierungsforschung, da viel darauf hindeutet, dass visuelle Informationen schneller verarbeitet werden können als Text. Wobei hier durchaus zu erwähnen ist, dass die tatsächliche Leistungssteigerung weitaus geringer sein dürfte als in den Medien oft behauptet wurde (z.B. '60.000 mal schneller als Text'). Außerdem deuten unsere Ergebnisse darauf hin, dass hinsichtlich der Fehlerquote bei der Beantwortung abschließender Fragen keine wesentlichen Unterschiede zwischen visuellen Informationen und Texten festgestellt werden können.

Abstract

As static visualizations like diagrams, maps, charts, or drawings get more and more important in everyday life, it is crucial to find out how helpful they actually are, especially in comparison to text. This diploma thesis outlines if visualizations or texts are processed faster by humans and which representation is better comprehensible. For this purpose, we conducted an exploratory study in which we measured processing times and error rates when interpreting either texts or visualizations.

In a pre-study, in which each participant had to describe four visualizations in their own words, we found out which parts of a visualization are deemed most important by people. The focus of the participants was on extrema, as well as on certain other values. Furthermore, the participants often compared different values in order to describe the visualization. The pre-study gave us valuable insights which we used for writing the texts for our study.

For our exploratory study, where we wanted to find out more about whether visualizations or text can be interpreted more easily, we used 15 visualizations and texts that contained the same information. Each participant had to work on at least six topics, thereof at least three by using visualizations and three by using texts, and had to answer three questions per topic at the end. The time was measured while the participants worked on the topics. We could see that the participants solved topics by using visualizations on average about 1.3 times faster as compared to when they used texts. This difference was statistically significant. We could not find significant differences between the error rates for topics, when participants used visualizations or texts. When texts were used, we found correlations between text lengths and processing speeds, as well as text lengths and error rates. The content of visualizations or texts did not seem to play a role for processing speed or error rates. However, we found cases in which topics with visualizations were solved more than 50 % faster as compared to topics with texts.

Our results provide a solid basis for defining further hypotheses regarding the readability of visualizations compared to text. In this thesis, we present the final hypotheses that emerge from our exploratory study. We consider these to be extremely interesting for visualization research, as there is much evidence that visual information can be processed faster than text. Whereby it is worth mentioning that the actual increase in performance may be much lower than often claimed in the media (e.g. '60,000 times faster than text').

Furthermore, our results indicate that no significant differences can be found between visual information and text with regard to the error rate in answering final questions.

Contents

CHAPTER 1

Introduction

Data visualization is a graphical representation of data or information. Data visualization uses visual elements such as charts, graphs, and maps, to provide the viewers with an easy and accessible way of understanding the represented information. With a visual representation, viewers can make sense of information by finding relationships in the data and confirm (or disapprove) ideas they had about the data. To convey ideas effectively, both aesthetics and functionality need to go hand in hand, which is discussed in data visualization design principles and guidelines. Data visualization has become an indispensable part of the business world and an ever increasing part of managing our daily life.

The history of data visualization goes back thousands of years. Early examples are maps, which are used by humans for about 8,000 years [Tho16]. The probably earliest town map that was found in Turkey dates back to about 6200 BC [FW21]. While tables are utilized at least since the second century, the idea to use graphical representations for displaying information started to form in the 17th century, when Rene Descartes created the foundation by inventing the two-dimensional coordinate system. In the 18th century, William Playfair invented many of the graph types that are still in use today, like, for example, bar graphs, line graphs, or pie charts. Jacques Bertin found out in the 1960s that visual perception is based on rules like, for example, the Gestalt Laws [Dah06], which can be used to display information visually such that it is understood better. In the 1980s, Edward Tufte [Tuf85] suggested that it is important to think about how data can be displayed effectively [Few14]. In the further course, and with the increase in computing power and the growth of the internet, the new discipline of information visualization evolved and advanced [CM97]. Nowadays, people come across visualizations multiple times a day. They can be found in many printed and online sources - on websites and social media, in newspapers, school books, or scientific journals. The benefits of data visualizations are advertised on several websites, which also suggest to apply these techniques as often as possible [Vis, Sib12, Ent18, Spi22]. Figure 1.1 shows different visualization types that have evolved over time [Bor13].

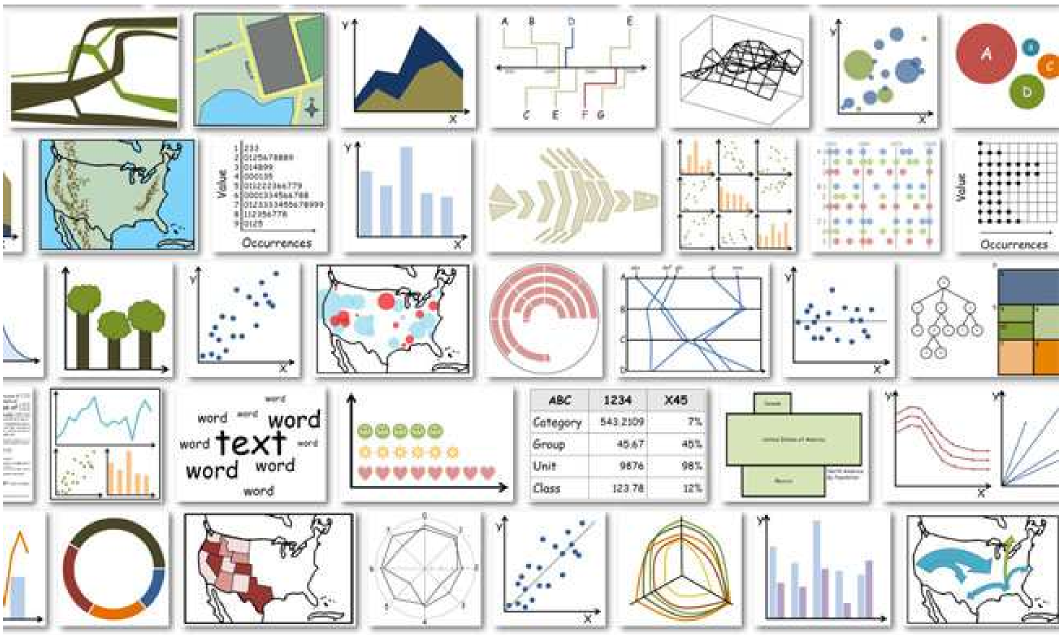


Figure 1.1: A graphic showing different visualization types that have evolved over time [Bor13].

1.1 Motivation

At the same time as data visualization was becoming more known and important, statements about the usage, the benefit, and the usefulness of visualizations have been made. One statement that can be found very often is that ‘the brain can identify images seen for as little as 13 ms’. This statement was confirmed in a study conducted by Potter et al. [PWHM14] and supports the claim that humans are very fast in interpreting visual information. For other statements, though, it is very hard to find confirmatory studies and references. For example, a very widespread claim that can often be found on websites or when advertising data visualizations, is: ‘Visuals are processed 60,000 times faster than text’. Although this would strongly favor the usage of data visualization, unfortunately, hardly any quantitative evaluation for this claim can be found. Many websites mentioning this claim often reference each other as sources, and some websites use citations from scientific publications as a justification [Sib12, Fot20, 3M97]. However, after an extensive research on the topic, it has to be stated that there seems to be no scientific source behind this statement. No studies have been found supporting this claim and it seems that this number (60,000 times) has never been measured. It is also unclear where this claim originates from. Some people tried to research the origin, but to date without success [Lev12].

Other statements which are not backed up by scientific studies are '90 % of information transmitted to the brain is visual', and that '65 % of the general population are visual learners'. The latter is mentioned in some scientific works, but without a detailed explanation where these numbers originate from. Similarly, only few quantitative information is available for the common saying 'a picture is worth a thousand words'. For example, it is not known how many words a picture or a visualization is actually worth and if there is a way to measure this. It is also unknown whether humans are better in processing visualizations or text and which one is faster and/or more accurate. Also, are visualizations or texts easier to understand? Shneiderman [Shn03] states that for some tasks, a visual presentation of the information is dramatically easier to use than a textual description. It is, though, unclear for which type of tasks this would apply.

1.2 Research Questions

While data visualization already has a long history, many of the probably exaggerated positive statements about the usefulness of data visualizations have been created within the last decades with the increased usage of computers. To be able to quantify these statements, it would be helpful to know if they can somehow be confirmed, and if they can somehow be quantified. This, in the further course, refers to the question how good, i.e., useful, data visualizations actually are, especially in comparison to texts. Finding answers to these questions becomes increasingly important as nowadays visualizations play a big role in our everyday life. In the further course answers to these questions will help in creating visualization guidelines and recommendations and will help towards a quantification of analysis results achieved by using data visualization.

To address this topic, we defined the following research questions to be the scope of this thesis:
Research Question 1: Do humans process visualizations or texts faster, and how much faster?
Research Question 2: Do humans comprehend information from visualizations or texts better, and how much better in terms of accuracy?
Research Question 3: Does the content play a role for processing speed or comprehensibility?

1.3 Methodology

To address the three research questions we applied the following methods:

1. **Literature research**: As a first step, the current state of the art on research and studies towards readability of data visualizations was analyzed. We wanted to find out if some of the research questions, or at least parts of them, have already been addressed in other areas.

2. **Pre-study**: The aim of this thesis is to compare visual representations of information with text only. In order to be able to write accurate texts for our study and to make sure the texts contain the same information as the visualizations, we conducted a pre-study where the participants had to describe visualizations in their own words and had to write a text so we could find out which are the most significant parts of a visualization.

3. **Exploratory study**: After getting an overview of related work, we decided to conduct an exploratory study to find answers to the named research questions. In our study participants were either shown visualizations or texts, and they had to answer three questions about the content they have seen/read. During this process, the time was measured to be able to calculate the processing speed. The correctness of the given answers was evaluated to measure accuracy and error rate. The study was conducted on a lab computer as a controlled user study, in which we observed the participants in order to be able to make sure they were not distracted during the study, or to intervene in case they had problems with certain tasks.

4. **Hypotheses**: As a result of our exploratory study, we formulated hypotheses for further research on this topic.

1.4 Thesis Structure

Chapter 2 starts with an introduction to information visualization and human perception, before discussing studies about visual attention, visualization literacy, processing speed, and comprehension of visualizations and texts. Chapter 3 describes how we chose the visualizations and text lengths for our exploratory study, as well as for our pre-study, before discussing the study design and results of our pre-study. In Chapter 4, our exploratory study is described, starting with goals, participants, and materials, before going over to introducing our task plan, the program we used to conduct the study, as well as how it was conducted. Also, the results of the study for comprehensibility and processing speed are discussed, how we evaluated them, as well as the results for the post study questions, content of the texts and visualizations, and the type of visualizations. Furthermore, some limitations of our study are discussed. In Chapter 5 we present a short summary of our work, conclusions, and future work.

CHAPTER 2

State of the Art and Related Work

This diploma thesis is situated in the field of information visualization. This chapter gives an introduction to the topics of information visualization, human perception, visual attention, and visualization literacy. Some studies are already available that measure processing speed and the comprehension of visualizations and texts. These studies and the results are also discussed in this chapter.

2.1 Information Visualization

Information visualization is the research topic of finding an effective mapping between data and visual form, and about creating visual images and interacting with them in order to solve a certain data-driven task. The purpose of information visualization is to amplify cognition, for example, by increasing memory and processing resources, enhancing the detection of patterns, or reducing the search for information [Car09]. Information visualization is distinguished from scientific visualization in a way that information visualization deals with abstract information, whereas scientific visualization concerns mostly physically-based and spatial data [Car09]. However, it can be difficult to determine if a visualization is an information visualization or a scientific visualization, because different definitions sometimes contradict each other and the boundaries between the two fields are blurry [TM04].

William Playfair is credited for inventing many of the graph types that are still in use today and helped, therefore, to create the foundation of information visualization [FW21, Few14]. One of Playfair's hand drawn visualizations that was published in one of his books [Pla86] in 1786 can be seen in Figure 2.1. It shows a line chart with England's imports (yellow line) and exports (red line) to/from Denmark and Norway in the course of 80 years. Since then, information visualization evolved, many more chart types were invented and visualizations are now mostly created with computers and often viewed on electronic devices, which also make it possible to interact with the visualizations.

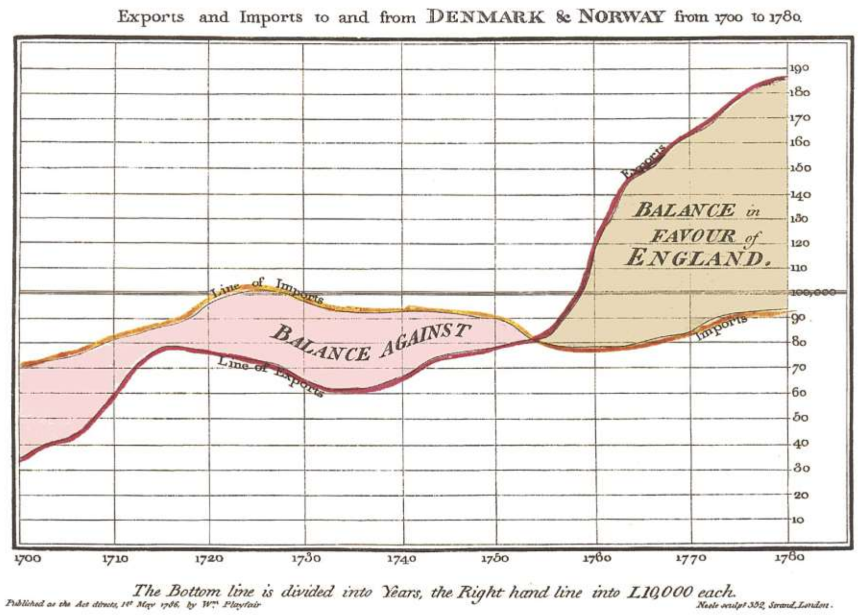


Figure 2.1: A visualization published by William Playfair in 1786 showing England's imports and exports to/from Denmark and Norway between 1700 and 1780. The yellow area denotes a positive balance for England, the red one a negative balance [Pla86].

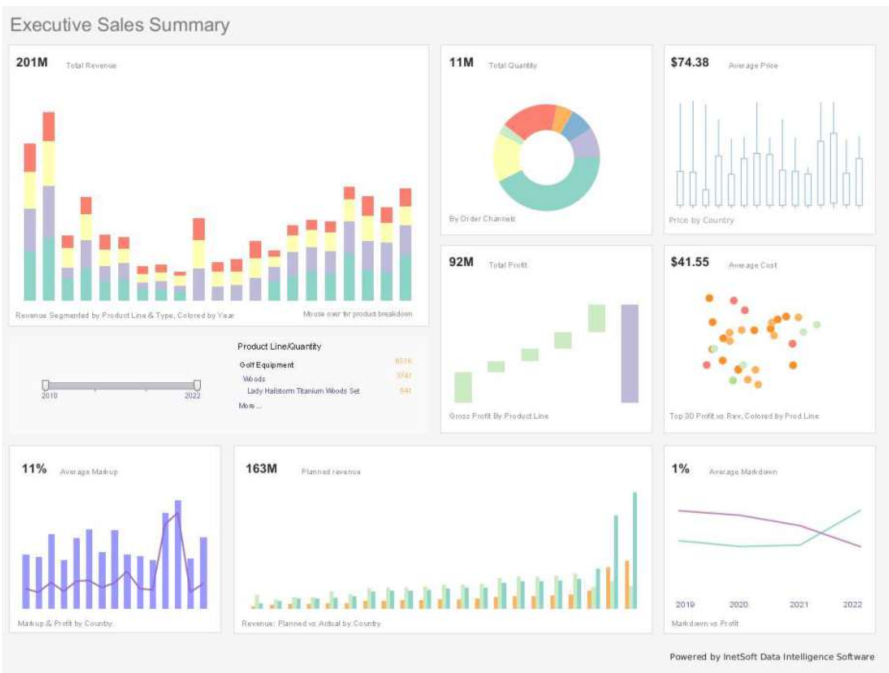


Figure 2.2: Modern information visualization dashboard showing an executive sales summary using different visualization types [ine].

A modern information visualization dashboard can be seen in Figure 2.2. Different visualization types are used to show an executive sales summary.

A good visual representation of a dataset can be helpful to solve tasks more effectively and can augment human capabilities [Mun14]. Visualizations are helpful for understanding models of complex phenomena. They help with analyzing and interpreting qualitative and quantitative information by providing fast and user-friendly mechanisms for data analysis [JMM+05, RTdOT06]. Visualizations also support decision making and can lead to new insights. A suitable information visualization can communicate and make humans comprehend a large amount of data in a short time. Due to this, visualization reveals patterns and connections that would be invisible otherwise. Visualizations can also show if there are any problems (e.g., errors or artifacts) in the underlying data [War19]. Card et al. [CMS99] stated that discovery, decision making, and explanation are the main goals of the insights gained through visualizations.

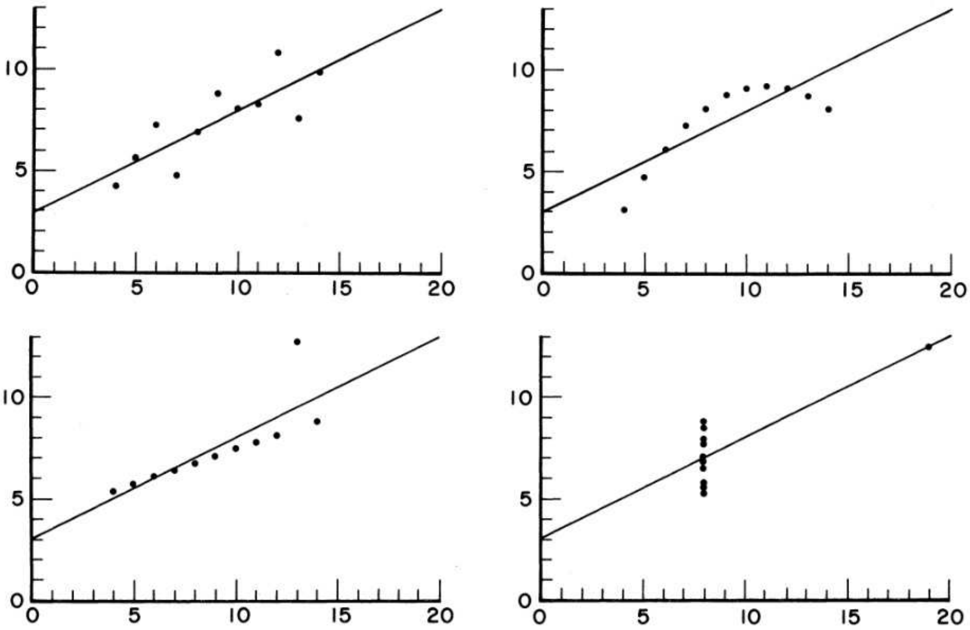


Figure 2.3: A visualization of Anscombe's Quartett. The graphs visualize four different datasets that share the same mean values for x and y, regression coefficients, and some other statistics. When visualized, they look completely different from one another [Ans73].

A famous example of a case where visualizations can help to get valuable insights into the underlying data is Anscombe's Quartett. Anscombe [Ans73] created four datasets with the same number of observations, mean values for x and y, regression coefficients, same equation of the regression line, and some other shared statistics. From a statistical point of view, without looking at the underlying data, the datasets would be considered to be equal. However, when visualized, the four datasets look completely different from one another. They can be seen in Figure 2.3.

According to Fekete et al. [FWSN08], information visualization is most useful for finding out which questions to ask about the data, or to find more meaningful questions about the currently inspected datasets. Especially more complicated questions, like, for example, if there is a correlation to be found between parameters or whether there are outliers in the data, can be answered more easily with visualization. It is also confirmed that visuals can help to memorize things better.

Van Wijk [VW05] argued that the amount of knowledge that is gained from a visualization also depends on the current knowledge of a person, because, for example, a specialist in a certain field might be able to extract more information from a certain visualization than a lay person. Further, the knowledge gain from a visualization depends on the underlying data and specifications, like hardware, algorithms, or parameters, that were used. Perception and cognition influence the knowledge gain as, for example, a color blind person might not be able to extract as much information as a person that can perceive all the colors in a visualization. There might be also differences in the abilities of people to spot patterns or structures in visualizations. The corresponding model can be seen in Figure 2.4. D is the data to be visualized, S the specifications of the visualization, V the visualization, I the image, P the perception of the user, K the knowledge of the user, and E the exploration done by the user. The increase in knowledge through the perception of image I is described by dK/dt, dS/dt describes the adaption of the specification of the visualization by the user based on the current knowledge in order to further explore the data.

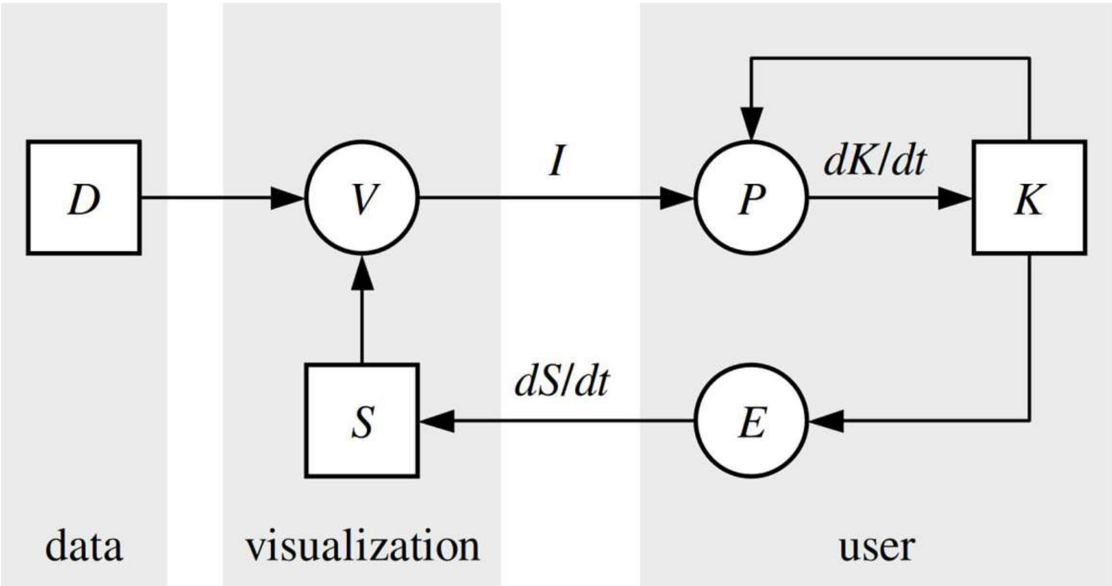


Figure 2.4: A model describing the relationships between data (D), visualization (V), specification (S), perception (P), knowledge (K), and exploration (E). I describes an image, dK/dt the increase in knowledge, and dS/dt the adaption of the specification by the user based on the current knowledge [VW05].

Creating a visualization means mapping data structures to visual structures [Maz09]. The data that is visualized can be divided into entities, which are the objects that shall be visualized, and relationships defining structures and patterns between entities. The purpose of a visualization may not only be to display relationships between entities, but also to find the relationships between entities. Both, entities and relationships, can also have attributes like colors. Attributes can have nominal, ordinal, interval, or ratio scales. A nominal scale is used for labeling, for example, different types of fruits. Ordinal scales are used for ordering things, for example, ranking films from best to worst. Interval scales are used when the gap between data values is of interest, for example, start and end time of a certain event. By using a ratio scale, the ratios between data values can be expressed, for example, that one value is twice as big as the other one. Attributes can also have one or multiple dimensions. A scalar might be used, for example, for the age of a person, while a vector might be used to display a traveling direction [War19]. Also, there are four elementary types of marks or graphical elements to display information in visualizations which are: points, lines, areas/surfaces, and volumes [CMS99, Maz09]. Graphical elements have properties like size, orientation, color, texture, and shape. Card and Mackinlay state that on top of marks and properties, visual presentations also consist of a position in time and space denoted by its coordinates [CM97]. Figure 2.5 shows the graphical elements on the left and properties on the right. Some of them are more effective than others, meaning that they can be interpreted faster and with less errors, are easier distinguishable than others, or better suited for representing certain types of data. Color, for example is effective for grouping data into different categories, but for displaying, for example, a sine wave, position should be used [Maz09, CMS99]. Cleveland and McGill [CM84] conducted a study about the accuracy of the perception of different graphical elements and properties. They were able to deduce a hierarchy and concluded

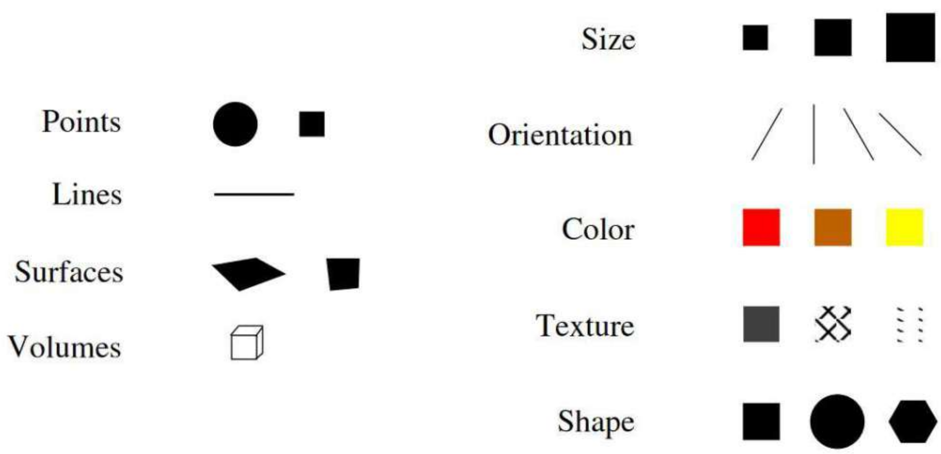


Figure 2.5: The four graphical elements on the left and some of their properties on the right. Some properties can be interpreted faster and with less errors, are easier distinguishable than others, or better suited for certain types of data [Maz09, CMS99].

that, for example, for displaying quantitative information orientation is perceived more accurately than color. As this is connected to how humans perceive visual content, we will discuss human perception in the next section.

Information visualization techniques have entered several different domains and application areas [CEH+08]. Whenever data of information that needs to be analyzed is very complex and heterogeneous, information visualization techniques have proven to be very helpful for users. This certainly applies to the medical domain, where understanding patients records [NYS09] is a very tedious task and where information visualization can also help to make patients understand their medical treatments better [WB16]. Climate research, where data of very long timespans has to be studied and understood, highly profits from the usage of information visualization [TCC22]. Again, communication of climate phenomena to lay users and the general public strongly depends on suitable visual representations [GMB16]. In Industry 4.0 applications, where a large amount of sensor data and other complex information needs to be analyzed, information visualization is considered as one of the key enabling technologies [AAJ+21]. Information visualization is used in sports to analyze games by displaying, for example, player positions and passes between the players [PVF13]. Visualization is used, for example, to monitor traffic situations, plan routes, or analyze the performance of public transportation systems [CGW15]. In finance, visualizations can help with monitoring large numbers of transactions or price fluctuations and decision making [DML14]. The contents and relationships of documents can be visualized to analyze word frequency, repetitions, topics, connections, or similarity between documents [GZL+14]. In marketing, information visualization can be used to analyze consumer behavior and spending patterns [YWLL20]. With the rise of data science as its own research field, the usage of information visualization techniques in different steps of the data analytics workflow became of increased interest for scientists [KPHH12]. This also led to more and more commercial applications being developed to support the visual exploration of data [BSS+18].

A more detailed analysis of all possible application fields of information visualization is beyond the scope of this thesis. The reader is referred to current data visualization conference proceedings and journals for an overview of current research results.

2.2 Perception

Perception is the process in which the world is interpreted and a mental representation of it is formed [WGK10]. Vision is the sense with the highest bandwidth, so, for us humans, most of the information is acquired through our visual channels, when compared to the other senses. The human visual system has evolved over millions of years by the interaction of the nervous system with the outside world and is the only channel used when interpreting visual information. Therefore, it is important to know how it works, in order to create suitable visualizations [War19].

Light is scattered from objects in the world and perceived by our eyes. A diagram showing the anatomy of the human eye can be seen in Figure 2.6. Six muscles are responsible for

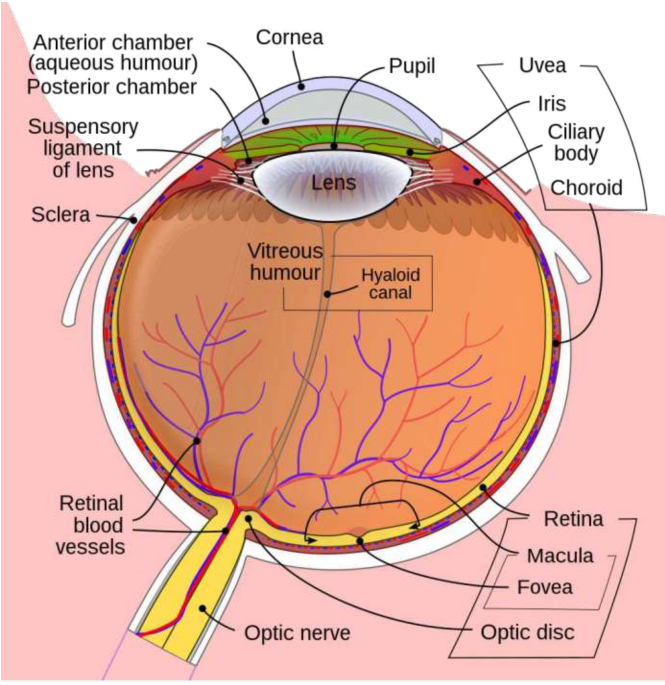


Figure 2.6: A diagram of the anatomy of the human eye [Rhc07]. Light is focused onto the lens by the cornea. The opening of the pupil determines the amount of light that enters the eye. On the back of the eye is the retina with its rods and cones. The fovea is a small region where vision is the sharpest. Visual information is sent to the brain via the optic nerve [WGK10].

moving the eyes to look at interesting parts of a scene. The cornea protects the eye from physical damage and focuses the light onto the lens. The pupil regulates the amount of light that enters the eye. Muscles change the curvature of the lens to be able to focus on objects that are near or far away. The retina is the photoreceptive layer onto which the light rays that enter the eye are projected. Rods and cones are the two types of photosensitive cells that make it possible to perceive intensity and colors. There are three types of cones for long, medium, and short wavelengths that correspond to red, green, and blue light. This can be seen in Figure 2.7. Humans can only perceive a small part of the spectrum between roughly 380 nm and 700 nm. The fovea is a small region, where vision is the sharpest, and it only contains cones [WGK10]. Only a small part of the visual field of the size of a thumbnail at arm's length is seen sharply and in high detail. This corresponds to a circular region with a diameter of 2 cm at a distance of 30 cm [BDDG03]. Images are acquired through cycles of fixations and saccades. During fixations, the eye stays at one position and gathers detailed information, while saccades are short, rapid movements between fixations [HE11]. Since the optic nerve only contains a limited number of fibers, the eye has to perform visual processing of the acquired data before it gets sent to the brain [WGK10].

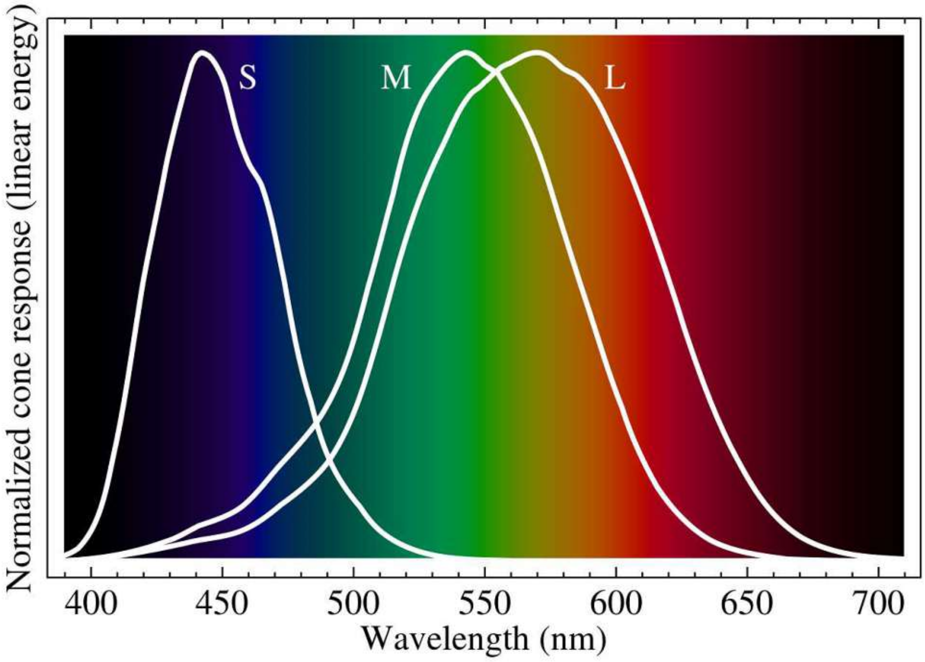
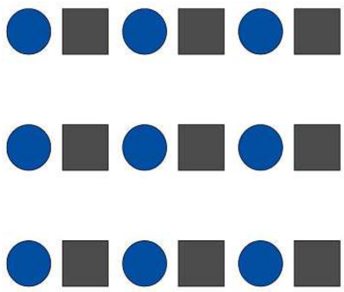


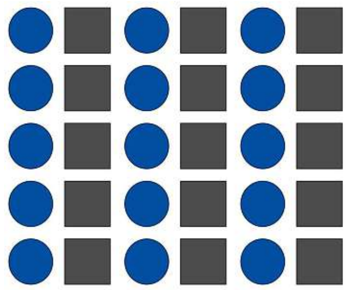
Figure 2.7: A diagram of the different wave lengths and color perception [Ben09]. The human visible spectrum is between 380 and 700 nm. Three types of cones are responsible for detecting short (S), medium (M), and long (L) wavelengths. These correspond to the colors blue (S), green (M), and red (L).

The brain makes sense of the data and stores it as memories. There are three types of memory: sensory memory, short-term memory, and long-term memory. The sensory memory gets impressions from sensor organs and retains them for a short period of time. It works automatically and independent of conscious control. The processing in sensory memory is called preattentive processing [Maz09]. Preattentive processing is automatic, performed in parallel, and is uncontrolled. There is also attentive processing, which is controlled, slower, and is used for short-term memory [WGK10]. Short-term-memory has limited capacity and information only stays there for a few seconds up to one minute unless it is rehearsed. When information is periodically rehearsed or meaningful associations are formed, it will be saved in long-term memory and can stay there for many years [Maz09].

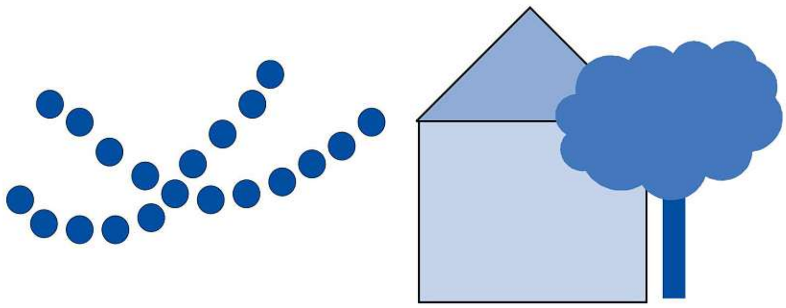
Ware [War19] proposed a perception model consisting of three stages. In the first stage, low-level features are extracted by rapid parallel processing of the scene. This process works unconsciously and cannot be controlled. In the second stage the features are processed in a slow and serial way, and the scene is parted in regions. In this stage also patterns are perceived. Working memory as well as long-term memory are involved in this second stage. In stage three, only a few objects, which are important for the current task, are held in the working memory. There are certain visual features that get detected very fast by visual processes. These features are called to be preattentive. Hue is such a preattentive feature, for example. A single red dot among many blue dots is spotted very easily and quickly. Other preattentive features are orientation, length, size, curvature, density, as well as direction and velocity of motion [HE11].



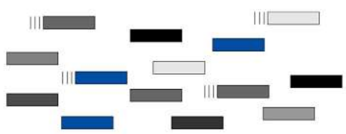
(a) Law of Proximity



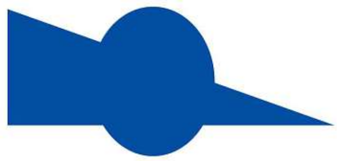
(b) Law of Similarity



(c) Two examples for Law of Good Continuation and Closure



(d) Law of Common Fate



(e) Example for both Law of Good Form and Law of Past Experience

Figure 2.8: Gestalt Laws [Dah06]. The Law of Proximity, which says that objects are perceived as one group when they are close together, can be seen in (a). (b) shows the Law of Similarity, where objects are perceived as one group when they look similar. (c) shows the Law of Good Continuation and Closure, which says that objects are seen as complete and also that we tend to follow lines. (d) shows an example for the Law of Common Fate, where objects that move in the same direction are visually grouped together. (e) displays the Law of Good Form and the Law of Past Experience. The Law of Good Form says that simple objects are favored over complex ones. The Law of Past Experience states that, if we learned that certain objects belong together, they will be perceived as belonging together in the future.

The Gestalt Laws by Wertheimer, Koffka, and Köhler describe different processes in human perception where preattentive features play a key role [Dah06]. Some very common Gestalt Laws are displayed in Figure 2.8. The first law described in this Figure is the Law of Proximity. It defines that objects which are placed close together are perceived as belonging to one group. This can be seen in Figure 2.8(a), where our eyes and brain merge the circles and squares together into horizontal lines. The Law of Similarity, as described by an example in Figure 2.8 (b), defines that we tend to group similar objects together. This is seen by the circles and squares which are merged into vertical lines. The Law of Good Continuation and Closure defines that objects are seen as complete by our eyes and brain, even if they are drawn incomplete or parts of an object are hidden. The law defines that we tend to follow lines and paths. Two examples for this law can be seen in Figure 2.8 (c). Humans will not focus on the gaps in the lines and will tend to follow the natural shape of the line. If two objects intersect, they are perceived as two complete objects, even if some parts are hidden behind the other object. The Law of Common fate defines that we perceive objects that move together in the same direction as one group. In Figure 2.8 (d) rectangles are shown, and for some of the rectangles movement is suggested by additional vertical lines. The 'moving' rectangles are perceived as one group, independent of the colors that are used here. Past experience also came to the conclusion that objects that were perceived as one group once, will also in the future always be perceived as one group. Figure 2.8 (e) shows an example of an object made out of two objects. If we learned in the past that these two objects frequently occur together, we will also perceive them as one in the future. This is known as the Law of Past Experience. Likewise, the Law of Good Form can also be described by Figure 2.8 (e), as it defines that objects with simple forms are favored by our perception over complex ones [Dah06].

Ware [War19] describes sensory and arbitrary symbols. The term 'sensory' is used for symbols and aspects of visualizations that can be processed by the brain without first learning them. They are detected by the visual sensory system and hard-wired processing and are understood without additional training. 'Arbitrary' refers to aspects that have to be learned first, for example, social aspects of the symbols in visualizations, as well as learned conventions. Sensory symbols have more expressive power when they are able to stimulate the visual system. Arbitrary symbols are better understood the better they were learned. An example for a hard-wired sensory phenomenon are optical illusions, which we still perceive 'wrong', even if we know the truth. This is because sensory codes happen unconsciously and cannot be controlled. Figure 2.9 shows an example for this, the Müller-Lyer Illusion. Even though we see that the middle parts of all arrows have the same length in the lower image, we still perceive them as if they had different lengths, when we look at the upper image. Another hard-wired process is the segmentation of the world into regions of the same color or texture. Also, sensory codes are normally understood across cultures. Contrary, arbitrary codes are constructed by society and have a meaning that is agreed upon. Arbitrary codes are hard to learn and easy to forget. They also depend on the culture, for example, different colors do not have the same meaning in different cultures.

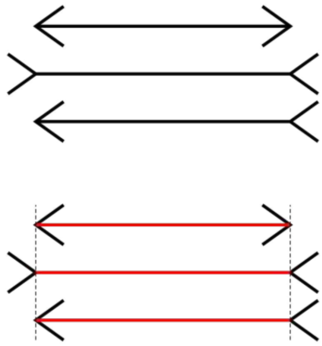


Figure 2.9: The Müller-Lyer Illusion, where the middle parts of the arrows are perceived like they had different lengths when actually they are equally long [Fib07].

Besides the Gestalt Laws, other rules and laws have been defined that describe human perception. Two of these other laws deal with perceived differences in visual representations. Weber's Law [Fec60] says that the likelihood to detect a change of a graphical attribute is proportional to its relative change, not the absolute one. This means, if an object is, for example, simply enlarged, it is not necessarily more effective at communicating information. Steven's Law [Ste57] states that the degree at which an attribute is underestimated increases as the attribute's dimensionality increases. This means that, for example, it is harder to estimate 3D volumes than 2D areas [WGK10].

The discussed aspects of perception all play an important role in how visualizations are perceived and should, therefore, be considered when designing visualizations. Only this way it can be ensured that data visualizations are as effective and comprehensible as possible and do not accidentally include misleading aspects [WGK10]. Different features (e.g., lines, rectangles, colors) can be used for data mapping (i.e., visually representing data attributes). However, a feature may draw more attention and may, therefore, be more salient than others depending on the task. For example, color is favored over shape by the human visual system. Consequently, the most important attribute in a visualization should be represented by the most salient feature [HE11].

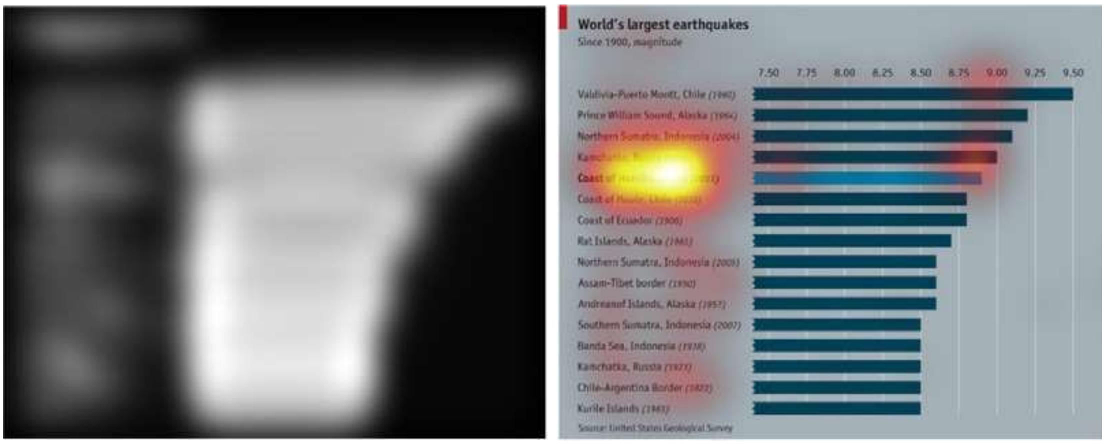
2.3 Visual Attention

The term 'visual attention' refers to the question which parts of an image or visualization get how much attention from the people who look at it. Some studies about this question have already been conducted. One of these studies was conducted by Borkin et al. [BBK+15]. In a large study they wanted to find out which are the parts of a visualization that people would remember after seeing it. Borkin et al. [BBK+15] used 393 different visualizations from the MASSVIS dataset [mas15, BVB+13, BBK+15]. Their 33 participants were shown 100 of these visualizations each, and every participant had to look at the visualization for 10 seconds. Eye-tracking was used to determine which parts of the visualizations were looked at by the people. Afterwards, the participants were shown 200 visualizations, each for 2 seconds, of which 100 were the visualizations previously shown to them. They had to indicate by pressing a button in case they recognized the visualization. Again, eye-tracking was used to record people's fixations on the visualizations. In the last part of the study, the participants were shown all visualizations they remembered correctly in the previous stage. However, this time the visualizations were presented at a smaller size and also blurred so that the participants could not read the text. Participants only were able to recognize which visualization it was. With this information presented, participants had to write down everything they remembered about these visualizations in 20 minutes. All in all 2,733 descriptions were written by the participants, with an average of 7.3 descriptions per visualization, and on average 13.1 words per description. The results of the study showed that 79,7 % of visualizations were recognized by the participants during the second stage of the study. The visualizations that were recognized, mostly contained human recognizable objects (e.g., company logos, flags, signs, pictograms, or icons). The elements in the visualizations that got the highest fixation times were titles and human recognizable objects. Concerning the descriptions, it was shown that the participants wrote the most, longest, and also the highest quality descriptions for such visualizations that were categorized as infographics. This implies that infographics were the most memorable types. The fewest descriptions and also shortest and of lowest quality were written for visualizations from governmental or scientific sources. Titles, labels, and paragraphs were the parts of visualizations that were most often written about in the descriptions. When titles were presented at the top of a visualization, they got more attention and got recalled more often than when they were at the bottom of a visualization.

Another study about visual attention was conducted by Polatsek et al. [PWV+18]. They took 30 visualizations from the news and infographics categories from the MASSVIS dataset. 16 of them contained human recognizable objects (e.g., pictograms or real objects like, for example, bottles). For every visualization, participants had to solve three different tasks. The participants either had to retrieve a value from the visualization, filter the data according to certain criterias, or find an extremum in the data. For each visualization, it was randomly chosen which one of these three tasks the participant would have to solve. In total, 47 participants completed the study. All participants were shown the task description first. When they understood it, they were shown the visualization without the task. After completing it, participants had to fill in answers into a form. The participants had to do this for all of the 30 visualizations. Eye-tracking data, task completion time, and correctness of the answers were recorded. The results showed that 14.1 % of all answers were incorrect. The task type with the most incorrect answers was the filter task. There was no significant difference in the correctness between the task types. The fixations collected during the experiment were compared to the fixations collected in the study by Borkin et al. [BBK+15]. It was shown that the participants fixated on similar parts of the visualization if they were given the same task type to solve. Also, the fixations of the tasks where participants had to find an extremum, were similar to the ones from the study by Borkin et al. [BBK+15]. The authors stated that extreme values in visualizations are often very conspicuous and get much attention, and, therefore, may be a key feature used by humans to remember the visualization. For all three tasks, the participants looked at the legend repeatedly while solving the task, while for the memorization task from the study by Borkin et al. [BBK+15] the legend was looked at only at the beginning. In general, text areas got more attention while people tried to remember the visualization, whereas data areas got more attention when people had to solve certain tasks. Participants took significantly more time to solve the filter tasks than to solve the tasks where they had to find an extremum or retrieve a certain value.

The property of a specific item to stand out in a visual scene is referred to as saliency. Saliency is often considered as a concept to describe important aspects in a data visualization. There are different saliency models that generate saliency maps of images or visualizations, so it can be seen which parts might get the most attention from the viewers. Matzen et al. [MHD+17] analyzed three of these models with 184 visualizations from the MASSVIS dataset, and compared the saliency maps to the fixations found in the study by Borkin et al. [BBK+15]. Matzen et al. [MHD+17] found out that all of these models performed worse on data visualizations than on natural images. Matzen et al. [MHD+17] concluded that existing saliency models are not adequate to predict the saliency of a data visualization. They also stated that text is not given enough importance by these models, even though textual parts of visualizations get much attention by the viewers. Also, bright red parts of visualizations were assigned a low saliency by the analyzed models, which contradicted actual fixations. Another problem of the analyzed models was that they focused on the center of the images, which is usually fine for natural images like photographs, but do not support the way data visualizations are built where the most important aspects are not necessarily placed in the center. Also, as opposed to natural images, visualizations may contain larger areas of the same color or white space, and the models might confuse them with each other. Therefore, Matzen et al. [MHD+17] proposed a new saliency model for visualizations, the Data Visualization Saliency Model (DVS). It is based on existing saliency models, but in addition uses text recognition, to put more attention to a visualization's textual parts. The DVS performed significantly better on data visualizations than the other existing saliency models.

Polatsek et al. [PWV+18] also compared eye-tracking data to saliency model results and DVS results. A saliency map generated with an existing saliency model and, in contrast to this, fixations during a task where participants had to retrieve a value from the visualization, can be seen in Figure 2.10. The whiter a region in the saliency map, the more salient it is. In Figure 2.10 (b) the white and yellow parts correspond to many fixations, red to a few. The light blue bar in Figure 2.10 (b) gets assigned a lower saliency in the map in Figure 2.10 (a) than the other dark blue bars. The authors found out that if people tried to memorize visualizations, DVS gave good predictions about where people would look. Contrary, if people had to solve a task, DVS performed significantly worse in predicting visual attention than for memorizing. Polatsek et al. [PWV+18] concluded that people focused more on the data parts of visualizations when they had to solve tasks, while DVS focuses more on textual parts, so it performed worse. For



(a) (b)

Figure 2.10: Two images taken from the study by Polatsek et al. [PWV+18]. (a) shows a saliency map generated with an existing saliency model. (b) shows recorded fixations during a task where participants had to retrieve a certain value. The whiter a region is, the more attention is on it.

memorizing visualizations, people also seemed to focus on the textual parts. Therefore, DVS performed well in this case.

2.4 Visualization Literacy

The term visualization literacy describes how well people can understand and also create visualizations (similar to literacy), and how well they can interpret the given information (e.g., discover patterns) [FJL22a]. So, even if a visualization is well designed, it does not necessarily mean that it is comprehended by all of the users [BRBF14]. Challenges in measuring visualization literacy include the huge number of different visualization types and the number of possible tasks that can be fulfilled with visualizations, or choosing a metric to measure visualization literacy [Ell18]. There might also be big differences in skill levels between participants and for different visualization types [SKND22]. Some other problems are that experts do not even agree on the names for some visualization types or on how to classify visualizations, e.g., as graph, chart, or diagram [BBG19].

Firat et al. conducted a detailed review on (interactive) visualization literacy to find trends and unsolved problems, and to give an overview of evaluation methods used in studies about visualization literacy [FJL22a]. They summarized their findings in a separate article [FJL22b]. Most of the reviewed related work used bar charts, followed by scatter plots and line charts. As a common future goal of most of the studies, Firat et al. [FJL22b] identified the understanding of barriers to visualization literacy by changing experimental settings, materials, or parameters of the studies. Many of the reviewed studies concluded that they would like to assess their approach with further visualization types or more participants/other participants, or that they plan to improve the used software, or improve the literacy test.

Boy et al. [BRBF14] created a series of tests with line graphs, bar charts, and scatter plots to measure visualization literacy in adults. They wanted to be able to filter out study participants with low visualization literacy. For the tests the authors used Item Respone Theory (IRT). IRT pays attention to the ability of a test person while also considering the item's difficulty. The probability for a test person to answer a question correctly is computed. An item consisted of stimulus, task, and question. A stimulus was the graphical representation, i.e., the visualization. Tasks included finding extrema, averages, intersections, making comparisons, etc. There were also different types of questions that were either about perception, e.g., retrieving the color of a certain element in the diagram, or about semantics, e.g., finding the highest values, or detecting whether there is a connection between different elements.

Lee et al. [LKK16] created a visualization literacy assessment test (VLAT). The test is especially suitable for non-experts in visualization and consists of 12 visualizations with 53 multiple-choice test items. For the visualizations, the authors chose the 12 most common visualization types which they deduced from a survey of news outlets and data visualization authoring tools. They included choropleth maps, bar charts, line charts, scatter plots, and area charts. The tasks included, among others, retrieving values, finding extrema, clusters, or anomalies, or determining a range. The tasks were presented as multiple-choice or true-false items. Classical test theory (CTT) was used to analyze the items. CTT pays attention to the difficulty of an item and if the item distinguishes between upper and lower scoring groups. The test was tried out with 191 test persons and showed a high reliability. The authors also found out that there was a fairly high correlation between visualization literacy and ability to learn an unfamiliar visualization.

Börner et al. [BBG19] proposed a framework to define, teach, and assess data visualization literacy. Their framework typology distinguished different insight needs, data scales, analyses, visualizations, graphic symbols, graphic variables, and interactions. Insight needs included clusters, comparisons, relationships, trends, and similar. The framework used the four scales nominal, ordinal, interval, and ratio, that were discussed in Section 2.1. Different types of analyses used to preprocess or analyze the dataset before visualizing it were considered in the typology. Included were statistical analyses, like, ordering or sorting the data, or relational analyses to examine relationships in the data. Six types of visualizations were distinguished in the framework typology: tables, charts, graphs, maps, trees, and networks. Different graphic symbols like points, lines, or volumes were considered, but also linguistic ones like texts, punctuation marks, and pictorial symbols like images or icons. Graphical variables included position, form, and color. Possible interactions were search, zoom, filter, extract, and similar. The typology can be taught to students and later used to assess visualization literacy, for example, by asking them to state the correct insight needs for a real world problem or classify a given visualization as graph, map, etc. The authors also defined a process model for the construction and interpretation of visualizations and linked the different stages to the typology.

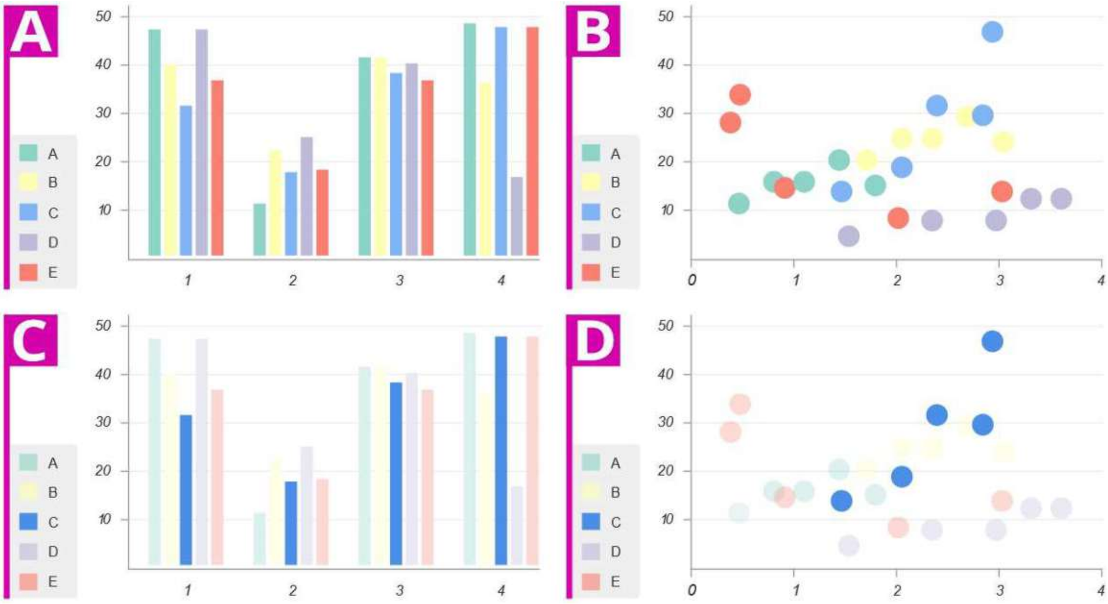


Figure 2.11: Examples for the visualizations used by Satkowski et al. [SKND22] to find out if adaptions have an influence in the performance of the participants depending on their level of visualization literacy. (A) and (B) show the original visualizations, (C) and (D) show the adapted forms in which certain colors were highlighted.

Satkowski et al. [SKND22] conducted a study about the influence of visualization literacy and adaptions of visualizations. Bar charts and scatter plots were either shown to the participants in their original form, or in an adapted version that highlighted certain bars or dots. Examples can be seen in Figure 2.11. (A) and (B) show the original visualizations, while (C) and (D) show the adapted ones. The study was conducted with 42 participants who had to answer multiple-choice questions about the visualizations. To achieve the correct results, participants had to filter, compute values, or determine a range. The visual literacy assessment of Boy et al. [BRBF14] was used to measure the participant's visual literacy prior to the study. Task completion time and accuracy were measured during the study. User experience measurements were determined by a post-study questionnaire. The questionnaire showed that participants slightly preferred the adapted visualizations over the original ones. Further results showed that there was an influence of visualization literacy on task completion time, where people with higher visualization literacy were faster in solving the tasks. No such effect was found for the accuracy. Adapted visualizations had, in general, a lower task completion time as compared to the original ones.

Mahmud et al. [MWW$^+$22] conducted a survey among 100 university students to find out how important visualization literacy skills are for them. The students had to submit ratings for statements from four different categories: the importance of visual literacy in learning, visualizations accompanying texts, additional visual learning in the curriculum, and communication with visualizations. A majority of students agreed or strongly agreed with the statements from all of the categories. This suggests that the students believed that visualization literacy contributes to learning, as well as when there are additional visualizations for a text. Their preferred learning approach was visual learning, and they believed that communicating information with visualizations was more efficient than without.

2.5 Processing Speed

In this section four studies and one comparative website article about processing speeds of visualizations or texts are presented. An overview can be seen in Table 2.1. The term 'processing speed' refers to the time a person needs to extract the information from a visualization or text, i.e., the time from the moment a person sees something until the information is extracted.

Study	**Number of Participants**	**Text (and Length)**	**Visualization**	**Condition**
Dunn [Dun19]	Comparative Website Article			
Thorpe et al. [TFM96]	15	-	Photographs	Vis.
Potter et al. [PWHM14]	32 + 32	-	Photographs	Vis.
Hauk et al. [HDF+06]	20	3 – 6 Letter Words	-	Text
Dyson and Haselgrove [DH00]	24	Articles (up to 1,000 Words)	-	Text

Table 2.1: Table of the studies about processing speed showing the number of participants, which texts were used, the text lengths in brackets, which visualizations were used, and the condition, i.e., visualization or text. Multiple numbers of participants in one cell denote the numbers of participants of different studies conducted in the respective paper.

Dunn [Dun19] aimed at finding better estimations for the claim that visuals are processed 60,000 times faster than text. They put together and compared the conclusions and results of different studies concerning processing times of written words and images. Among the compared studies are those by Thorpe et al. [TFM96], Potter et al. [PWHM14], and Hauk et al. [HDF+06]. Based on the results of these studies, Dunn estimated image processing times to be between 13 ms and 150 ms, the time to process a sentence containing 25 words to be between 3.75 seconds and 7.5 seconds, and the time to process a sentence with 8 words between 1.2 seconds and 2.4 seconds. Utilizing these numbers, Dunn draws the conclusion that images are processed between 8 and 577 times faster than words and rounds it to between 6 and 600 times.

Thorpe et al. [TFM96] conducted a study where the processing speed of images was measured. Between 700 and 2,000 photos were shown to the participants who had to judge if the photo contained an animal or not by pressing or releasing a button. The results indicate that even if a photo was only shown for 20 ms, the participants were able to categorize 94 % of the images correctly (on average).

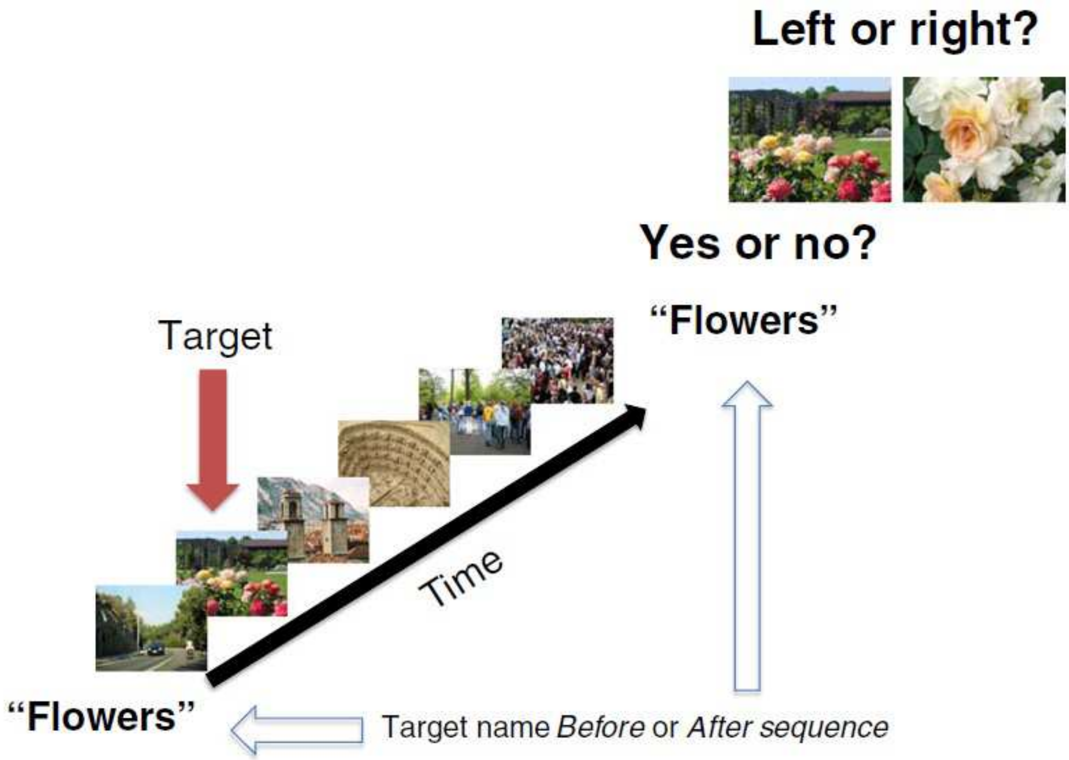


Figure 2.12: Illustration showing how the participants had to detect a target picture in a series of six pictures in the study by Potter et al. [PWHM14]. The study participants were asked to decide if a target picture specified by a name (e.g., 'flowers') could be found in the series of pictures. Then they also had to specify which picture exactly was shown.

Another study about the processing speed of images was conducted by Potter et al. [PWHM14]. A series of six or twelve pictures, in which each picture was presented for 13, 27, 53 or 80 ms, was shown to the participants, who had to detect if a target picture specified by a name, e.g., 'campfire', was present in the stream of pictures. The target name was either shown before or after the stream was shown. If the stream contained the picture, the participants had to identify this image out of two images with the same target name. The process is illustrated in Figure 2.12. The detection rate was always better when the target name was shown before the stream, independent of the showing durations and if the stream contained six or twelve pictures. The only exception was the six picture stream with a showing duration of 13 ms. In this case, the detection rate was better when the target name was shown after the participants looked at the stream. Participants were able to detect targets with a likelihood of more than 50 %, even if they only saw them for 13 ms; the longer they saw them, the higher became the detection rate, though.

Hauk et al. [HDF+06] conducted an experiment about visual word recognition speed. Participants had to decide by pressing a button if a word was an actual English noun or an invented word. The words were between three and six characters long. Each was shown for 100 ms. The results establish that participants were significantly faster for actual words than for invented words and also the error rate was higher for invented words.

Dyson and Haselgrove [DH00] investigated the effects of reading speed on reading comprehension. In an initial training phase, participants had to read a text in twice their normal reading speed. For the subsequent study three documents, containing up to 1,000 words, had to be read in normal speed and three at a faster speed while the time was measured. After each document the participants answered multiple-choice questions about the content, and specified if certain sentences appeared in the text or not. On average, in the normal reading speed, 4.06 words were read per second, compared to the faster one with 7.66 words per second. The comprehension scores as well as recall of certain sentences were better at the normal reading speed in all cases.

To summarize, user studies so far concentrated on understanding specific parts of the information processing pipeline with images and texts. As a cumulative evaluation of different studies, Dunn [Dun19] compared processing times from different studies. However, it has to be noted that these studies were conducted under different conditions and with different goals and hypotheses. Consequently, only a rough estimate can be made for the differences between processing times of visualizations and texts. Only photographs have been used in the studies (no data visualizations) [PWHM14, TFM96]. Hence, measurements for further and different visualization types like graphs or diagrams are needed. Also, it has to be pointed out that in some of the studies [PWHM14, DH00] the participants were trained to achieve high processing speeds, for example, by being asked to read twice as fast, or shortening viewing intervals for photos further and further. The performance and comprehension of the contents were always better when the participants were able to process the materials for a longer time period.

2.6 Comprehension

In this section 18 studies and two meta-analyses about the comprehension of visualizations and texts are presented. An overview can be seen in Table 2.2. These studies measured the participant's comprehension, i.e., how well and correct information is understood by test persons in case content is presented as either text or visualization.

2.6.1 Judgement of Comprehension

Serra and Dunlosky [SD10] conducted two experiments. In the first one, students had to study a text about lightning storms, either accompanied by diagrams, or not. The participants had to judge how good they think they will comprehend the materials before they began, after each paragraph, and after they were finished with studying.

Study	**Number of Participants**	**Text (and Length)**	**Visualization**	**Condition**
Serra and Dunlosky [SD10]	80 + 120	Science Text (ca. 500 Words)	Photograph or Diagram	Text vs. Text + Vis.
Ikeda et al. [IKT+13]	60 + 46	Science Texts (7 Paragraphs x ca. 300 Letters)	Image or Bar Graph	Text vs. Text + Vis.
Levie and Lentz [LL82]	Meta-Analysis			
Guo et al. [GZWM20]	Meta-Analysis			
Chibana [Chi19]	1,428	Articles (250 Words)	Infographics	Text vs. Vis.
Petrova and Riekhakaynen[PR19]	22	Science Texts	Infographics	Text vs. Vis.
McCrudden et al. [MSLP07]	47 + 55	Science Text (1,385 Words)	Diagram	Text vs. Vis., Text vs. Text + Vis.
Branch and Riordan [BR00]	274	Science Text (3 Pages)	Diagram	Text vs. Vis.
Erfani [Erf12]	65	English for Special Purposes Texts	Pictures	Text vs. Text + Vis.
Pan and Pan [PP09]	95	Text about Traffic Accident (123 or 162 Words)	Drawings	Text vs. Text + Vis.
Ardasheva et al. [AWR+18]	174	Science Texts (424 – 444 Words)	Photograph or Drawing	Text vs. Text + Vis.
Holmqvist et al. [HOWBN17]	46	Art Texts (75 - 91 Words)	Paintings	Text vs. Text + Vis.
Reinking et al. [RHM88]	167	Science Texts (3 x 250 - 300 Words)	Diagram, Map, or Illustration	Text vs. Text + Vis.
Roberts et al. [RNC15]	156	30 Words, 7 Passages and 19 Sentences	Diagram, Table, Chart, etc.	Text + Vis. (unrelated)
Jian and Ko [JK17]	42	Science Texts (400 + 414 Chinese Characters)	Photograph or Diagram	Text + Vis.
Firat [Fır17]	99	Maths Text Questions	Charts and Pictures	Text vs. Text + Vis.
Berends and van Lieshout [BvL09]	135	24 Maths Text Questions	Drawings	Text + Vis.
Schnotz et al. [SLU+14]	40	Science Texts (36 – 116 Words)	Photograph, Diagram, Map, or Graph	Text + Vis.
Lenzner et al. [LSM13]	30 + 57 + 194	Science Texts (1,130 Words)	Photograph or Diagram	Text vs. Text + Vis.
Yarbrough [Yar19]	89	Management Texts	Infographics	Text vs. Text + Vis.

Table 2.2: Table of the studies about comprehensibility showing the number of participants, which texts were used, the text lengths in brackets, which visualizations were used, and the condition, i.e., visualization, text or both. Multiple numbers of participants in one cell denote the numbers of participants of different studies conducted in the respective paper.

Subsequently, they had to answer comprehension questions. The results show that the students thought that additional diagrams would lead to better comprehension. An analysis of the latencies indicated that more time was spent on the text with diagrams compared to text-only, but additional diagrams also led to a better performance when answering the comprehension questions. In their second experiment, a third group was added, which saw images of lightning strikes instead of diagrams. Examples for the images and visualizations can be seen in Figure 2.13. The group with images estimated their comprehension to be even a bit better than the one with diagrams did. However, the best performing group was the one that used text with diagrams, while the text-only group performed worst. Also, the total studying time was the highest for the group with images.

(a) (b)

Figure 2.13: A diagram showing how lightning storms develop (a) and an image of lightning strikes (b), used in the study by Serra et al. [SD10]. Both images were accompanied by texts.

The study by Ikeda et al. [IKT$^+$13] investigated their participants' judgement of comprehension before and after reading a text. In a first experiment, the participants saw either only a text or text accompanied by brain images, and then they had to answer comprehension questions. The results confirmed that the text accompanied by images was judged to be better comprehensible than the one without images, but there was no significant difference between the two groups. In a second experiment, texts with brain images or bar graphs were used. Examples can be seen in Figure 2.14. The text with images was believed to be better comprehensible, but again, there was no difference in actual comprehension. Therefore, the authors concluded that judgements about comprehension conflict actual comprehension, and that in this study images and bar graphs did not help the participants to understand the text.

To summarize, both studies found that people judge visualizations to increase comprehension, but their expected positive effects seem inflated compared to the actual measured results.

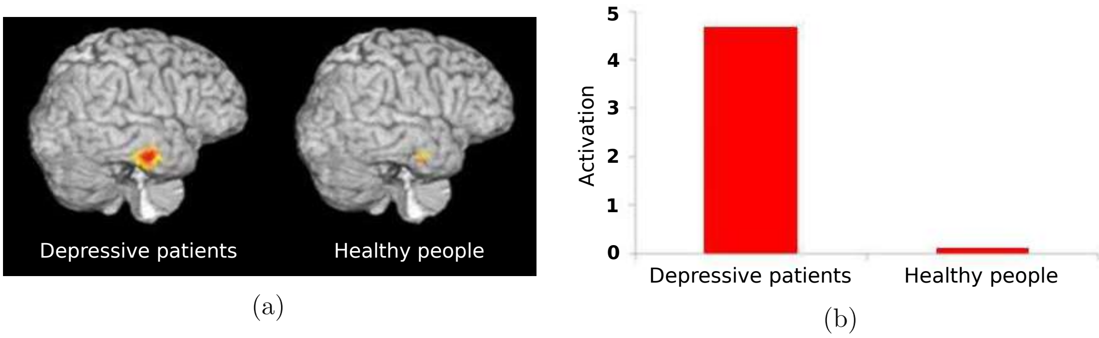


Figure 2.14: Examples of brain images (a) and bar graphs (b) that were used by Ikeda et al. [IKT+13]. Both images were accompanied by texts and show the same information in different ways.

2.6.2 Meta-Analyses

Meta-analyses compare and summarize results of different studies about a topic and use statistics to form conclusions.

Levie and Lentz [LL82] did a review and summary of 55 studies conducted between 1938 and 1981, that examined if illustrated text helps with learning as opposed to text alone. They did not consider studies using graphs, charts, or diagrams. Most of the studies in their report used either photographs or drawings as illustrations. 46 comparisons from 23 studies about learning information from text and illustrations showed that 45 reported better mean scores for the groups that saw illustrated text compared to the groups that saw text alone. In 39 cases the differences were statistically significant. On average, the scores of the groups with illustrations were better by 36 %. In another comparison of 10 studies, examining if illustrations have detrimental effects when they are only decorative, it was concluded that illustrations have no effect in this case. In 48 comparisons that did not fit into the groups with the other comparisons, 38 favored illustrated texts.

Guo et al. [GZWM20] did a meta-analysis about the impact of graphics on reading comprehension where they analyzed 39 studies conducted between 1985 and 2018. Most of them compared the usage of text-only with text presented together with graphics. Just two studies investigated the impact of text-only versus graphics-only. In the meta-analysis, it was concluded that graphics overall had a moderate positive effect on reading comprehension. Also, the comprehension improved if pictures were added to the text as opposed to combinations of different graphic types. Comparing the effects of different graphic types to each other, all of them, i.e., pictures, pictorial diagrams, and flow diagrams, showed similar beneficial effects on comprehension. No significant differences between grade levels of the participants were found. However, the authors state that something to be considered is, that the graphics for the studies were specifically crafted to be helpful for learning the texts, and were also of high quality and created by experts. Therefore, the positive effects that were found may be inflated, and it is hard to say if they are universal and applicable for visualizations in general.

Both of the discussed meta-analyses came to the conclusion that illustrations have at least moderate positive effects on the comprehension of the content.

2.6.3 Text and additional Visualizations

In a vast majority of comprehension studies text was compared to text again, but with additional visualizations to illustrate the content. In these studies researchers analyzed the effects of different cueing conditions, as well as impacts for English learners, adults with dyslexia, and people from different age groups.

Erfani [Erf12] conducted a study in which Iranian students were divided into two groups, both were taught for twelve weeks either by using only English texts or English texts with additional pictures. Afterwards, they had to do a test. The results show that pictures enhanced the students' reading comprehension. The instructor noticed that students who were taught using pictures participated more in discussions, and were more attentive.

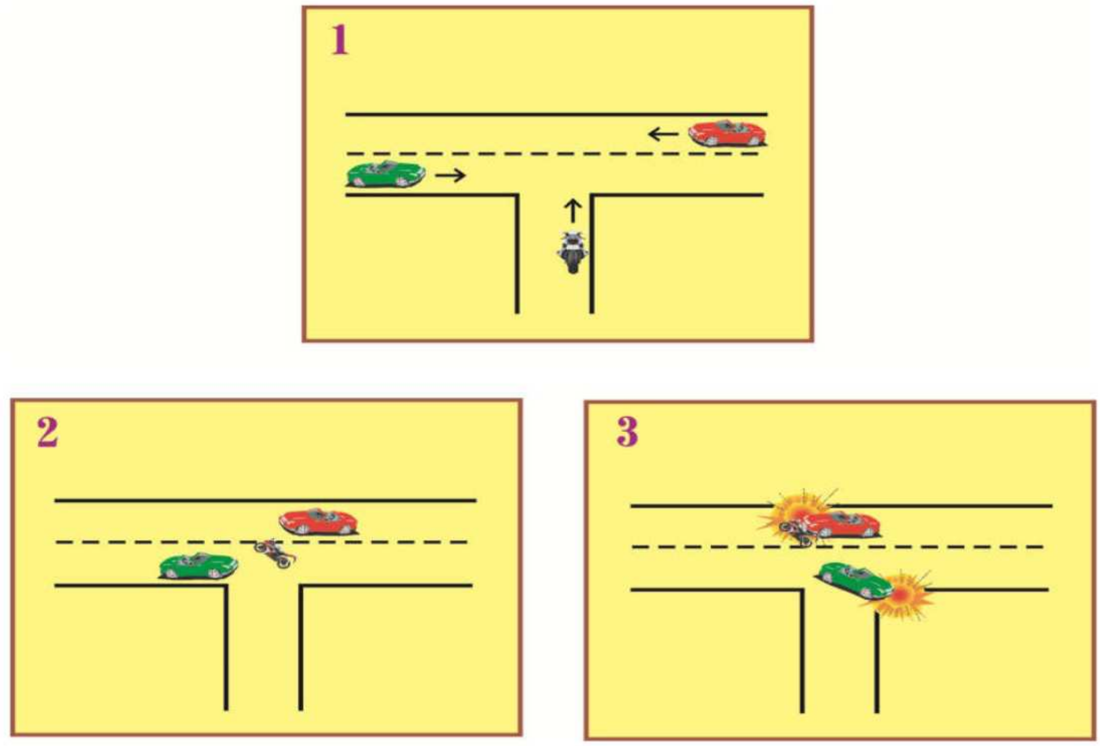


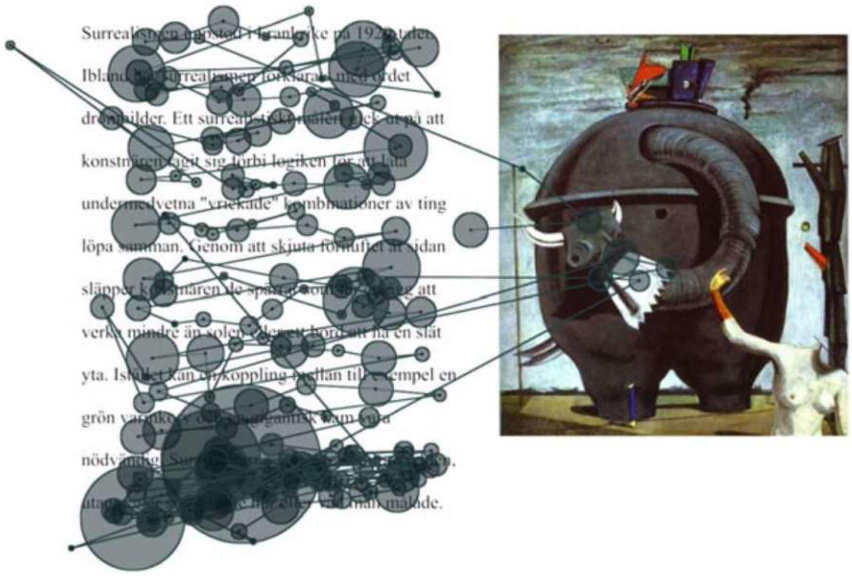
Figure 2.15: Drawings describing an accident, as used in the study by Pan and Pan [PP09]. The drawings were shown to the participants in addition to a text that had to be translated.

The effects of pictures alongside with texts for low proficiency English learners were examined by Pan and Pan [PP09]. Their participants were Taiwanese college students. Each of them had to read an English text describing a traffic accident, either on a lower or higher difficulty level, and with or without drawings visualizing what happened in the accident. The drawings can be seen in Figure 2.15. The students had to translate the English text into Chinese and rate the pictures' helpfulness. It can be seen in the results that the translation scores were significantly higher with additional pictures for both difficulty levels. The students overall agreed that pictures helped them understand the text better, but to different extents when comparing texts with lower and higher difficulty level. Students reading the text with lower difficulty level found illustrations to be more helpful for guessing the meanings of unknown words in the text, while students reading the text with higher difficulty level found illustrations to be more helpful for understanding what happened.

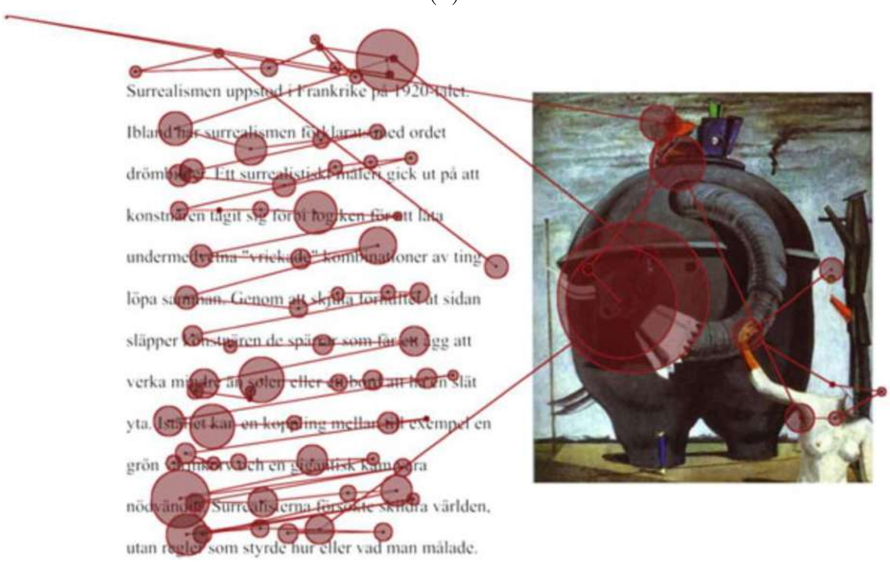
The results of a study with seventh-grade English learners by Ardasheva et al. [AWR+18] were different than those of Erfani [Erf12] and Pan and Pan [PP09]. Two physics texts had to be read by the participants, which were either accompanied by photographs and drawings, or not. The drawings and photos were labeled and visualized the contents of the texts. After reading the texts, the students had to answer comprehension questions and questions about their interest in physics. Concerning the comprehension scores and triggered interest, the results of both groups showed no significant differences, but a small tendency favoring the text without visuals. The authors state that no positive effects of the utilized visuals in this study were found.

In an eye-tracking study by Holmqvist et al. [HOWBN17] the effects of illustrations on reading patterns and text comprehension of adults with dyslexia were measured. People with dyslexia have problems with reading and understanding words or texts, even though they have no impaired vision or hearing. The participants were assigned to one of two groups, which either only saw texts, or texts accompanied by a picture. The same study was also done with the control group. Six texts about different art genres had to be read by each participant, the group with pictures also saw an example painting of the respective genre next to the text. When the participants decided they had studied the content long enough to answer questions about it, they had to click a button and were then asked what they remembered about the genre, and their answers were recorded. Afterwards, they were also given multiple choice questions. Examples for the recorded scan paths can be seen in Figure 2.16. On average, people spent 49.1 seconds viewing the screen containing text without pictures and 50 seconds with pictures. When pictures were present, people with dyslexia spent more time looking at the text compared to text without pictures. For the control group, it was the opposite. Overall, only 2.2 seconds were spent looking at the pictures as opposed to 45.9 seconds for looking at the texts. Concerning the learning effect of pictures, people with dyslexia scored worse when a picture was presented with the text.

A study with seventh and eighth graders by Reinking et al. [RHM88] examined the effects of different cuing conditions for graphics. The authors wanted to find out if there is a difference in comprehension of the visualizations if people were told to look at the provided graphic, or explicitly at certain parts of a graphic. The materials for the study were a text and two types of graphics of which one contained redundant information



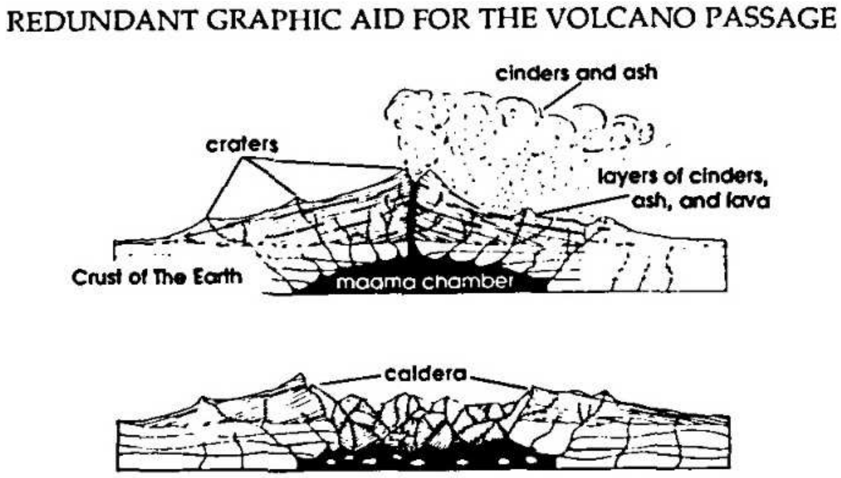
(a)



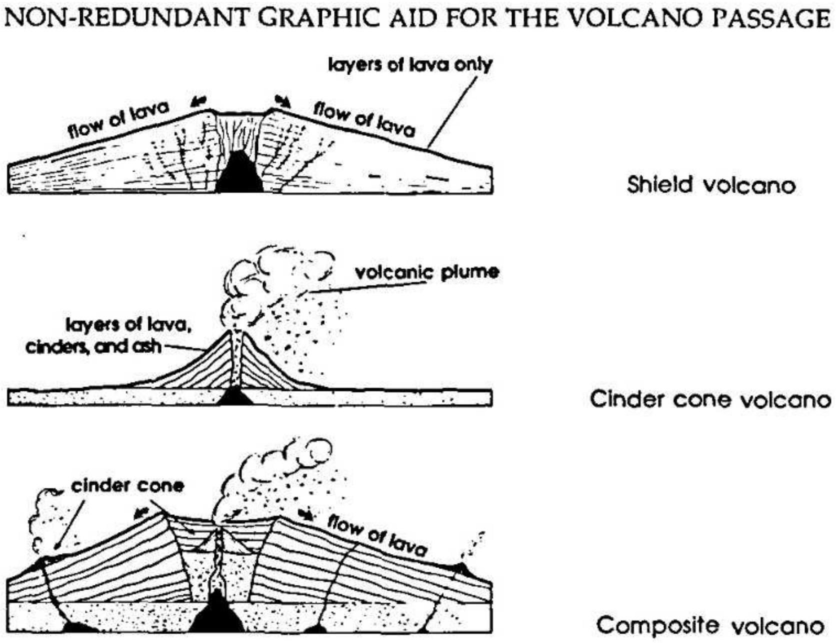
(b)

Figure 2.16: Examples for scanpaths of a person with dyslexia (a) and a person from the control group (b) from the study by Holmqvist et al. [HOWBN17]. When a picture was presented in addition to the text, people with dyslexia spent more time looking at the text compared to without pictures. It was the opposite for the control group.

and the other one information that was new, but related to the text. Examples can be seen in Figure 2.17. There were also four different cuing conditions for the graphics. The different cues were used to refer the readers to certain graphics or certain parts



(a)



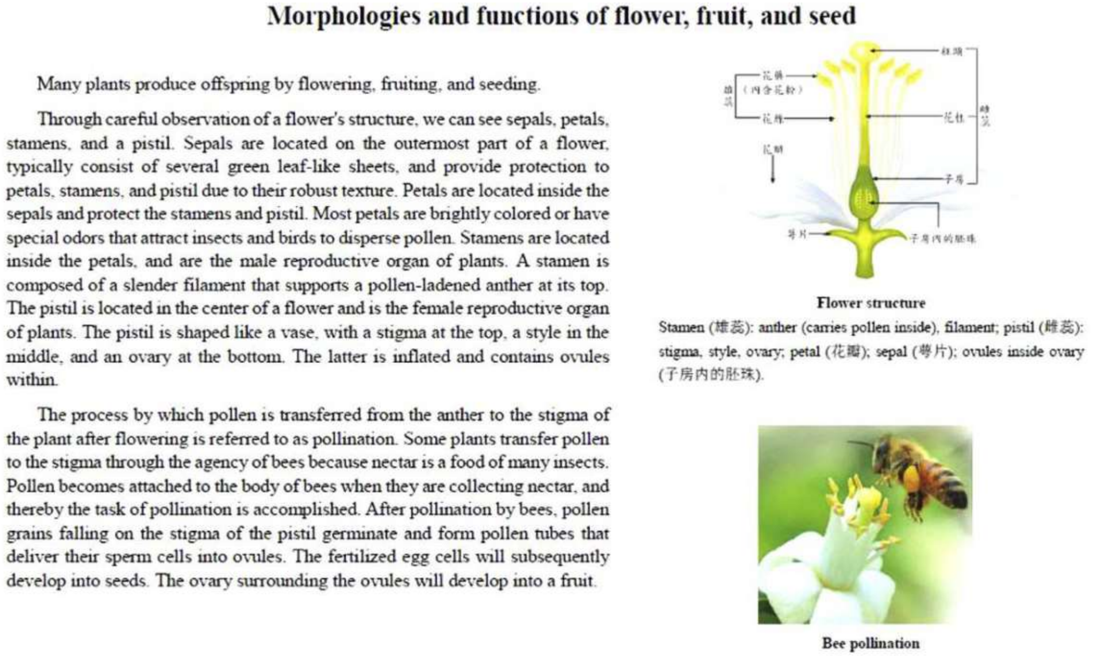
(b)

Figure 2.17: Examples for the visualizations used by Reinking et al. [RHM88]. The diagrams in (a) contained information that was also written in the text, while the diagrams in (b) showed additional information that was related to the text.

of the graphics. The conditions were: without cues, general cues, specific cues, and combined cues. Together with another condition, where only the text without graphics was used, the students were separated into five groups. General cues only told the reader to examine the graphics for better understanding, while specific cues directed the attention to specific parts of the graphics. Concerning the graphics, for example, diagrams, maps, or illustrations were used. After studying the materials, the participants had to fill out multiple choice tests in which they had to recall what they learned, and also label different parts of the graphics. The results show that specific cuing increased the attention for, and recall of, graphic aids. Furthermore, subjects with one of the cuing conditions, especially those with the specific and combined cuing conditions, could better recall information than those without cues. More questions about the redundant graphics were answered correctly compared to the related graphics with new information. Also the difference between good and poor students' performances was less for graphics with any form of cuing, but cuing in general also benefited good readers.

Figure 2.18: Example of an illustration the children in the study by Roberts et al. [RNC15] had to understand. For every illustration, children had to describe what is shown in the illustration.

A study by Roberts et al. [RNC15] examined the reading comprehension of children as well as their comprehension of different graphical devices like pictures, flowcharts, diagrams, or tables. Third graders were asked to name and explain seven different graphical devices. They also had to complete other tasks like gaining information from the graphics, and answering questions, or identifying a certain type of graphic in a set with other graphics. An example for an illustration can be seen in Figure 2.18, where the children had to describe what is shown in the illustration. The study authors also did a reading assessment where the children had to identify a word in a group of other words or had to find synonyms. In another assessment, children had to read passages aloud, and it was measured how many words they read correctly per minute and how well they comprehend the text. From the results, the authors concluded that reading comprehension correlates with graphical device comprehension. It also depends on the childrens' development how well they are able to understand different visualizations. Even if the students are from the same grade, there might be great variations in understanding.


Morphologies and functions of flower, fruit, and seed

Many plants produce offspring by flowering, fruiting, and seeding.

Through careful observation of a flower's structure, we can see sepals, petals, stamens, and a pistil. Sepals are located on the outermost part of a flower, typically consist of several green leaf-like sheets, and provide protection to petals, stamens, and pistil due to their robust texture. Petals are located inside the sepals and protect the stamens and pistil. Most petals are brightly colored or have special odors that attract insects and birds to disperse pollen. Stamens are located inside the petals, and are the male reproductive organ of plants. A stamen is composed of a slender filament that supports a pollen-ladened anther at its top. The pistil is located in the center of a flower and is the female reproductive organ of plants. The pistil is shaped like a vase, with a stigma at the top, a style in the middle, and an ovary at the bottom. The latter is inflated and contains ovules within.

The process by which pollen is transferred from the anther to the stigma of the plant after flowering is referred to as pollination. Some plants transfer pollen to the stigma through the agency of bees because nectar is a food of many insects. Pollen becomes attached to the body of bees when they are collecting nectar, and thereby the task of pollination is accomplished. After pollination by bees, pollen grains falling on the stigma of the pistil germinate and form pollen tubes that deliver their sperm cells into ovules. The fertilized egg cells will subsequently develop into seeds. The ovary surrounding the ovules will develop into a fruit.

Flower structure

Stamen (雄蕊): anther (carries pollen inside), filament; pistil (雌蕊): stigma, style, ovary; petal (花瓣); sepal (萼片); ovules inside ovary (子房内的胚珠).

Bee pollination

Figure 2.19: One of the illustrated biology texts used by Jian and Ko [JK17]. The children had to read the illustrated text, and had to answer questions about it afterwards.

Jian and Ko [JK17] did an eye-tracking study with 10-year-old children about reading comprehension of illustrated biology texts. Figure 2.19 shows one of the illustrated text examples. The children were separated into two groups depending on their reading abilities. Afterwards, they had to read the same two illustrated science texts, having different complexities, and answer comprehension questions. The children with higher reading ability performed better for all questions, and both groups did better on the text of lower difficulty. While both took roughly the same amount of time for reading, the high-ability group rather focused on the difficult text and pictures, while the low-ability group was more focused on the text in general, and especially the one of lower difficulty. The authors concluded that children with lower reading ability had problems understanding the illustrations and linking them to the text.

Firat [Fır17] analyzed the effects of visuals on the test performance of elementary students from fifth grade. The students had to solve one of three different forms of maths tests, either with only text, text and 'model' visuals, or text and 'real' visuals. 'Model' visuals were, for example, diagrams or tables, while 'real' visuals contained, for example, images

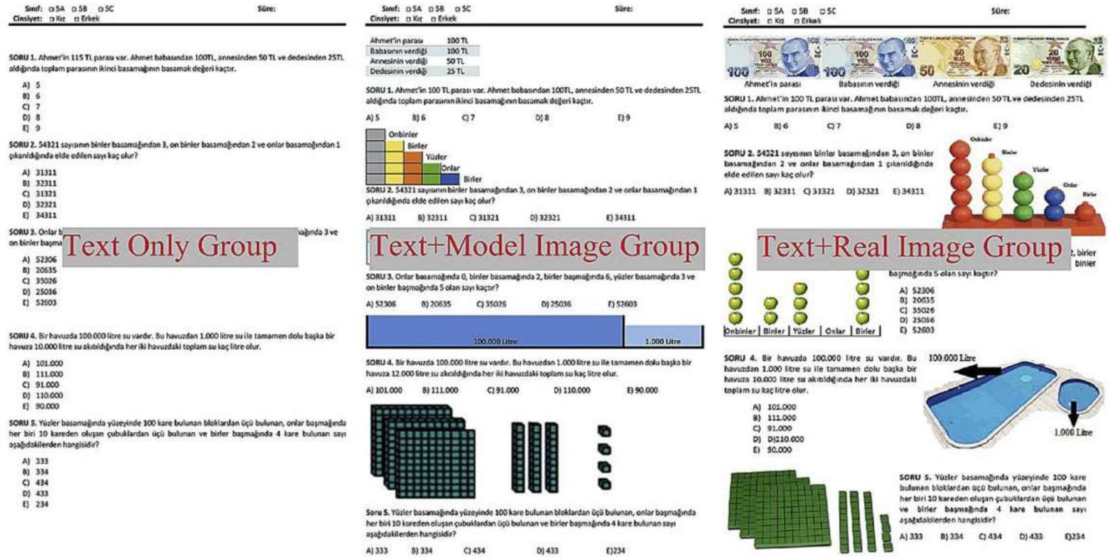


Figure 2.20: The three types of maths tests from the study by Firat [Fır17]. The image on the left shows the test that only contained text and no illustrations. In the middle image is the test with text and 'model' visuals like diagrams or tables. The right image shows the test with text and 'real' visuals like bank notes or swimming pools. Each student had to solve one of the three different tests.

of bank notes or swimming pools. The three types of tests can be seen in Figure 2.20. The authors found out that students who did tests with visuals had higher scores than those who only saw text questions. Also, the average time needed to complete the test was the highest for the test with 'model' visuals and the lowest for the test with 'real' visuals.

Berends and van Lieshout [BvL09] conducted a study with fifth graders and maths questions in which the speed and accuracy of solving arithmetic word problems (i.e., maths text questions) presented with four different types of illustrations were measured. The illustration types were 'bare' (e.g., equations for the word problems), 'useless' (i.e., decorative illustrations), 'helpful' (i.e., illustrations representing what was written in the word problem), and 'essential' (i.e., illustrations that contained additional information). Examples are shown in Figure 2.21. The participants saw the same 24 arithmetic word problems accompanied by the different illustration types. Dependent on their score in a standardized maths test, they were assigned to the good or poor arithmetician group. Results show that the scores of both groups were lower and the time they needed was higher for the word problems with essential illustrations compared to the other ones. Word problems with 'bare' illustrations were solved fastest. The authors state that using illustrations together with word problems does not necessarily improve the performance, but may even have a negative impact, especially if the presented illustrations are irrelevant for solving the problem.

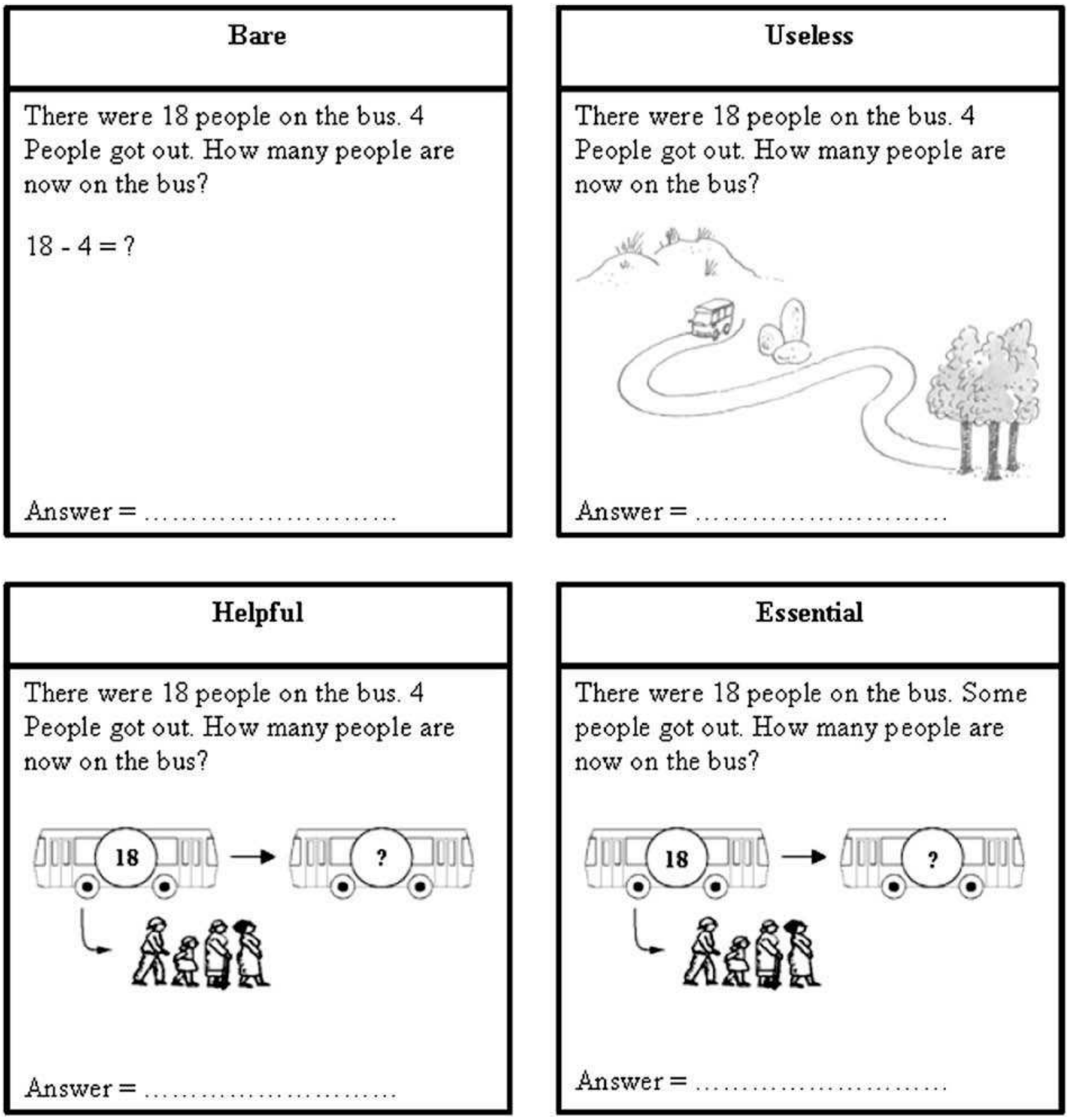


Figure 2.21: Examples of the different illustration types used in the study by Berends and van Lieshout [BvL09]. The type 'bare' did not contain any illustration, but just the equation. 'Useless' illustrations did not help to solve the math problem and were only decorative. 'Helpful' illustrations showed exactly what was asked for in the math question. 'Essential' illustrations contained information that was not written in the question but had to be understood in order to answer the question correctly.

In an eye-tracking study by Schnotz et al. [SLU+14] the authors examined students' strategies to integrate information from text and pictures. The participants were from grades five and eight from the higher and lower tier of the German school system. Four

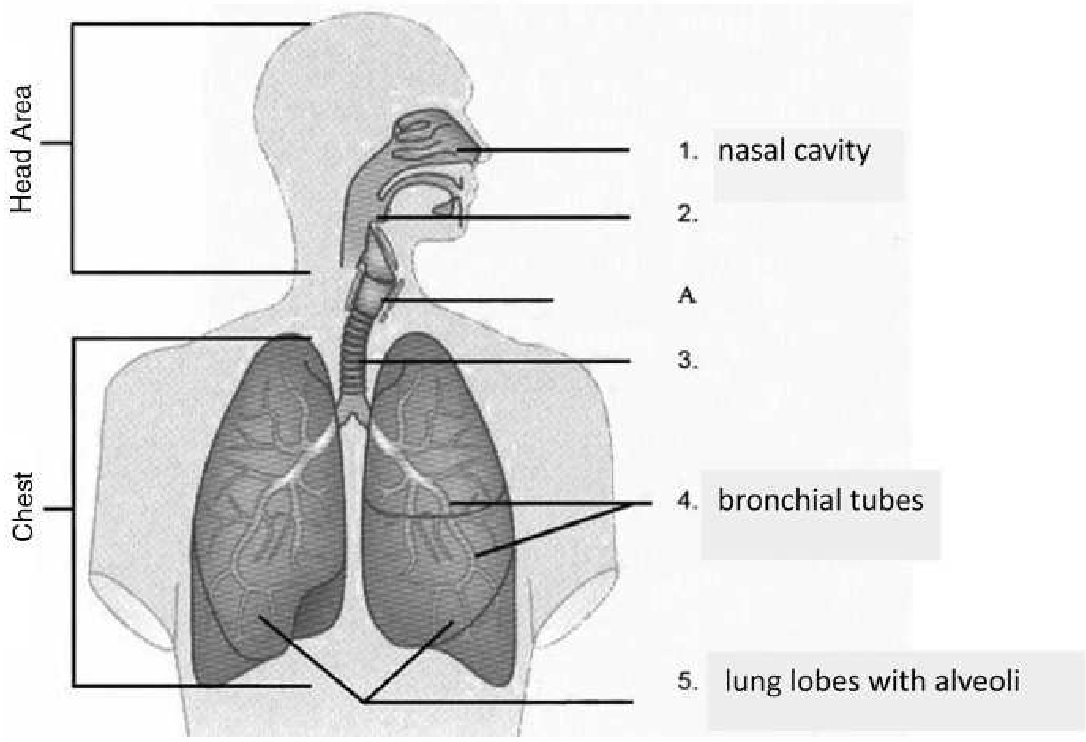


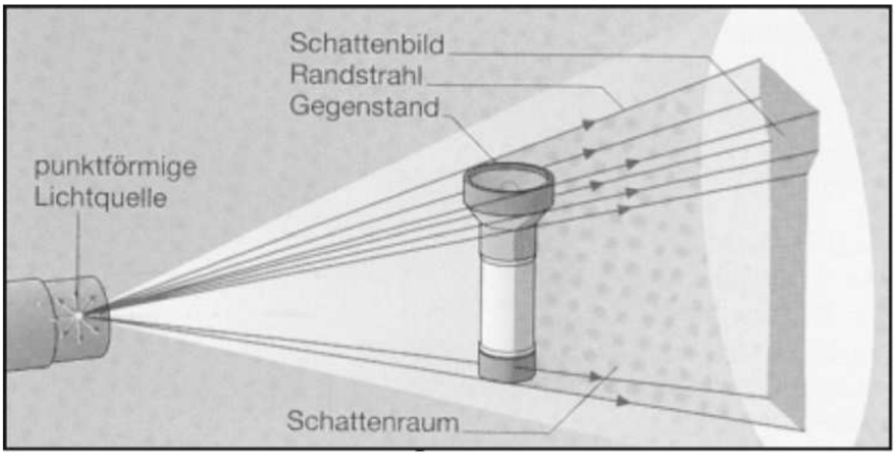
Figure A3. Respiratory system: The respiratory system ensures that air enters the body when inhaling and that it leaves the body when exhaling. As shown in the diagram, the air passes through the respiratory system, which is labeled with numbers representing the specific respiratory organs. Number 2 represents the pharynx. The trachea (No. 3) connects the respiratory organs in the head area with the respiratory organs in the chest. To prevent choking, the epiglottis (A), which is not part of the respiratory organs, covers the entrance of the trachea. In the chest, the trachea bifurcates into two main bronchial tubes, which in turn branch off, ending in millions of air-filled alveoli. All of the alveoli together make up the two lungs lobes.

Figure 2.22: A diagram as visualization and image caption as textual part, used in the study by Schnotz et al. [SLU+14]. The diagram contained information that was not written in the caption, and the caption contained information that was not written in the diagram. In order to be able to answer the questions, the participants had to understand the diagram, as well as the caption.

texts with additional pictures that illustrated information not mentioned in the texts were presented to the students together with three comprehension questions with growing difficulty. To answer the questions, the text as well as the pictures and how they are linked had to be understood. In most cases, the texts, that were used together with the images, were image captions for the visualizations. An example for a diagram and related text (i.e., the diagram's caption) can be seen in Figure 2.22. The authors concluded that texts are used for global understanding, while pictures are used for selective processing like finding the relevant information for a task. Concerning the measured processing speeds, the average time for processing text, pictures, and the questions, was 652 seconds.

Students from the higher tier gave more correct answers and needed less time for reading than students from the lower tier. The text fixation times were always highest when answering the first question, even though it was the easiest one, and the lowest for the last question, the most difficult one. Concerning fixations on pictures, it was the same effect for students from the lower tier. Students from the higher tier had the lowest fixation times on the picture during answering the second question. The highest fixation times for students from the higher tier were reached during answering the third question.

(a)



(b)

Figure 2.23: Example of a decorative picture (a) and instructional picture (b) about shadows used by Lenzner et al. [LSM13]. The participants either saw a text with both pictures, with only one of them, or without a picture.

The role of decorative and instructional pictures in learning was examined in three experiments by Lenzner et al. [LSM13], which were conducted with seventh and eighth graders. In a first eye-tracking experiment the students got presented either a text with decorative or instructional pictures, or both. For a text about shadows, for example, the instructional picture was a diagram illustrating how a shadow forms, while the decorative picture showed trees throwing shadows. Both pictures can be seen in Figure 2.23. It was found that decorative pictures were only given attention at the beginning and were afterwards ignored, while instructional pictures were viewed throughout reading the whole text. Instructional pictures were also viewed two to three times longer compared to decorative ones. In the second experiment the same material was used to study whether decorative pictures improved the students' mood and lead to more alertness and calmness compared to instructional pictures. The results showed that decorative pictures have indeed these hypothesized positive effects. In the last experiment an additional group of participants was included that only got to see a text without pictures. In this experiment it could be seen that students perceived the learning material to be less difficult if decorative pictures were present and that they were more interested when they had instructional pictures. Also, the learning performance was better with instructional pictures, while decorative pictures were found neither to be harmful nor helpful for learning. However, combined with decorative pictures, the instructional pictures lead to an even better learning performance.

Yarbrough [Yar19] analyzed if infographics improve online learning by using them for weekly summaries of the content of an online course. Every week students could do optional quizzes about the infographics. One of the infographics is shown in Figure 2.24. At the end of the course students were asked about their opinions on infographics and especially regarding learning with them. The study showed that a majority of students found them helpful for summarizing and remembering key concepts, and when learning for the final exam. Compared to the same section of the course without infographics, the final course grades for the section with infographics were better on average.

In general, the studies using text with additional visualizations mostly found positive effects for recall and comprehension. One study showed that references to visualizations in the text help people to understand visualizations and draw attention to them, especially when certain parts of the visualization are referenced explicitly [RHM88]. However, also negative effects of visualizations were mentioned. Some of the studies concluded, that particularly younger children or poor learners can have difficulties in integrating information from text and visualizations and understanding the correlations [JK17, BvL09, SLU$^+$14, RNC15]. Another study also found that additional visualizations can have negative effects for people with dyslexia [HOWBN17]. Concerning these studies, it may be hard to measure the effects of visualizations alone, because the additional text also contributes to comprehensibility, and it is especially hard to tell how well a visualization without text would be understood. The fact that the attention would have to be split between text and visualizations is also the reason why these studies are not viable for comparing processing times of visualizations to those of texts.

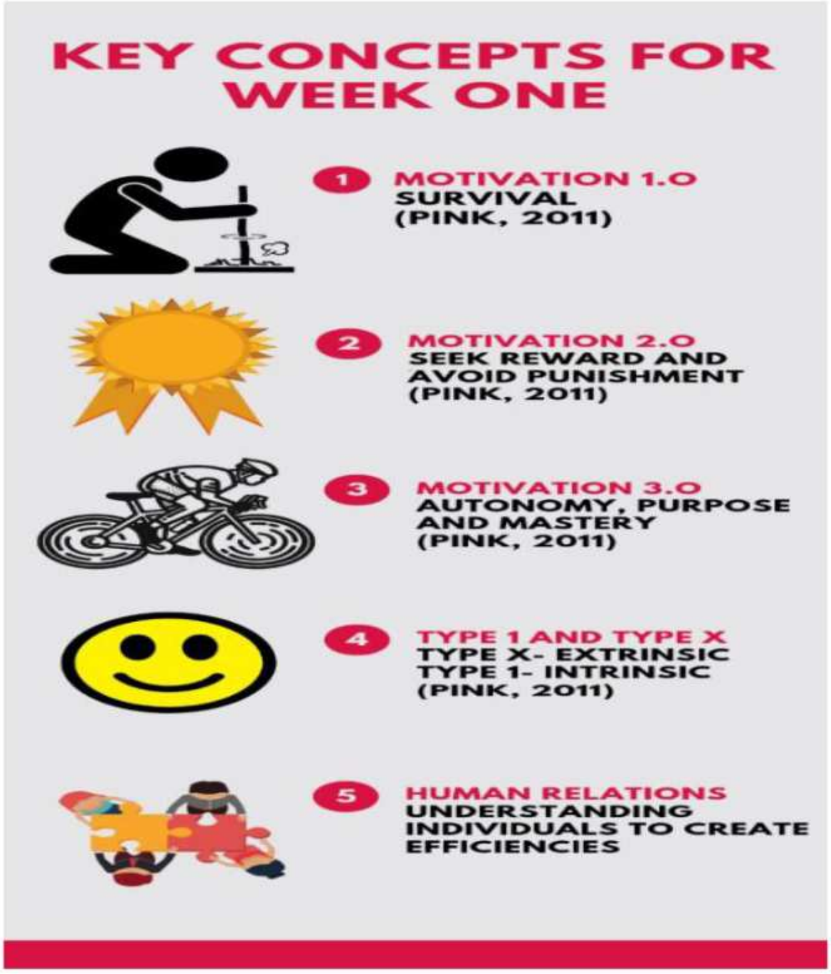


Figure 2.24: One of the summarizing infographics used in the study by Yarbrough [Yar19]. Each week during an online course, the students got a summarizing infographic with the key concepts and could answer optional questions about it.

Two studies analyzing the effects of visualizations for English learners concluded that they were helpful [Erf12, PP09], while one found, that the visualizations had no positive effects at all, and the study even had a small tendency favoring texts [AWR+18].

Also, especially concerning decorative visualizations, the studies showed conflictive results. It was found that they can divert the attention from the important things and time may be spent unnecessarily to process the visualization [BvL09]. On the contrary, also positive effects, like an increase in alertness, calmness, and learning performance were found [LSM13].

2.6.4 Visualizations-only versus Text-only

Most of the studies that compared the comprehensibilty of visualizations as opposed to text compared texts to texts with different additional visualizations, while only a few compared text to visualizations without text. These studies are especially interesting and important as they directly refer to the topic of this diploma thesis. The following four studies directly compare text-only information to visualizations-only.

Chibana [Chi19] published a study where each of the 1,428 participants were presented a total of six articles. The articles were either presented as a text containing 250 words, or as an infographic with the same information. Afterwards, the participants had to answer comprehension questions. The results show that depending on the type of infographic, participants sometimes scored higher and sometimes lower on the comprehension questions than participants that had to read a text. The same accounts for their recall of the material. However, in most cases, differences were insignificantly small.

Petrova and Riekhakaynen [PR19] conducted an eye-tracking study with 22 students between 14 and 17 years comparing infographics to text, and also measured participants' processing times. Four infographics with corresponding texts were used in the study. All students saw two different infographics and two different texts, and had to answer questions about the content and difficulty afterwards. For three of the four assignments participants gave more correct answers, and the difficulty was estimated to be lower for the infographics than their textual counterparts. However, the questions concerning general understanding were answered more accurately in three of four cases when text was provided. The authors stated that text is better suited for reaching conclusions, while infographics are better suited for remembering and understanding local parts. Concerning the total processing time, there was only a significant difference in one case where the text was long and, therefore, more time was spent on the text as opposed to the corresponding infographic. It was concluded that the processing time rather depends on text characteristics like topic, size, and design. Furthermore, text areas in infographics tended to be observed longer than picture areas, of which some had no fixations at all.

McCrudden et al. [MSLP07] studied the effects of causal diagrams on text learning. Causal diagrams show which causes lead to which effects by linking them with arrows. In their study they used a diagram depicting the effects of space travel on the human body, which can be seen in Figure 2.25. In a first experiment the participants had to read and memorize the contents of a text in ten minutes, either with or without a diagram, and answer retention and comprehension questions afterwards. There were no significant differences in memory or comprehension concerning the main ideas of the text, but the causal sequences were understood better with a diagram. Also, the more complex a sequence, the more helpful the diagram. The second experiment was conducted with diagram-only versus text-only conditions using the same materials as in the first experiment. Even though one group only had a diagram and no text, there were no significant differences in comprehension between the groups. Therefore, it was concluded that a well-designed diagram can contain as much information as the corresponding text.

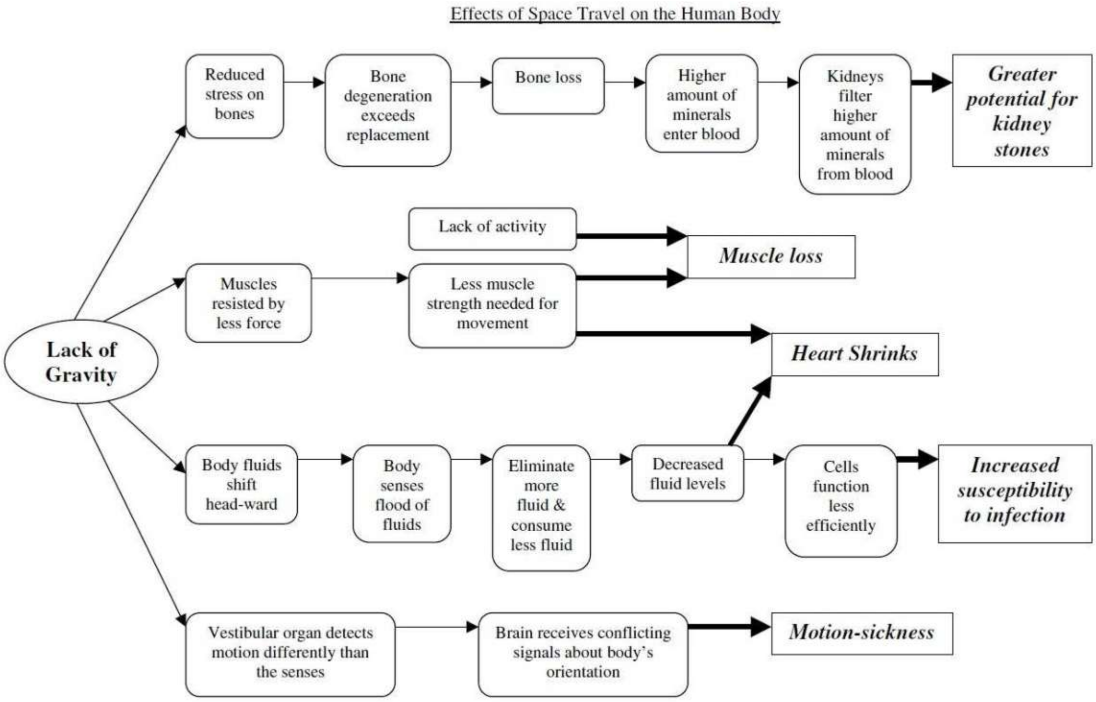
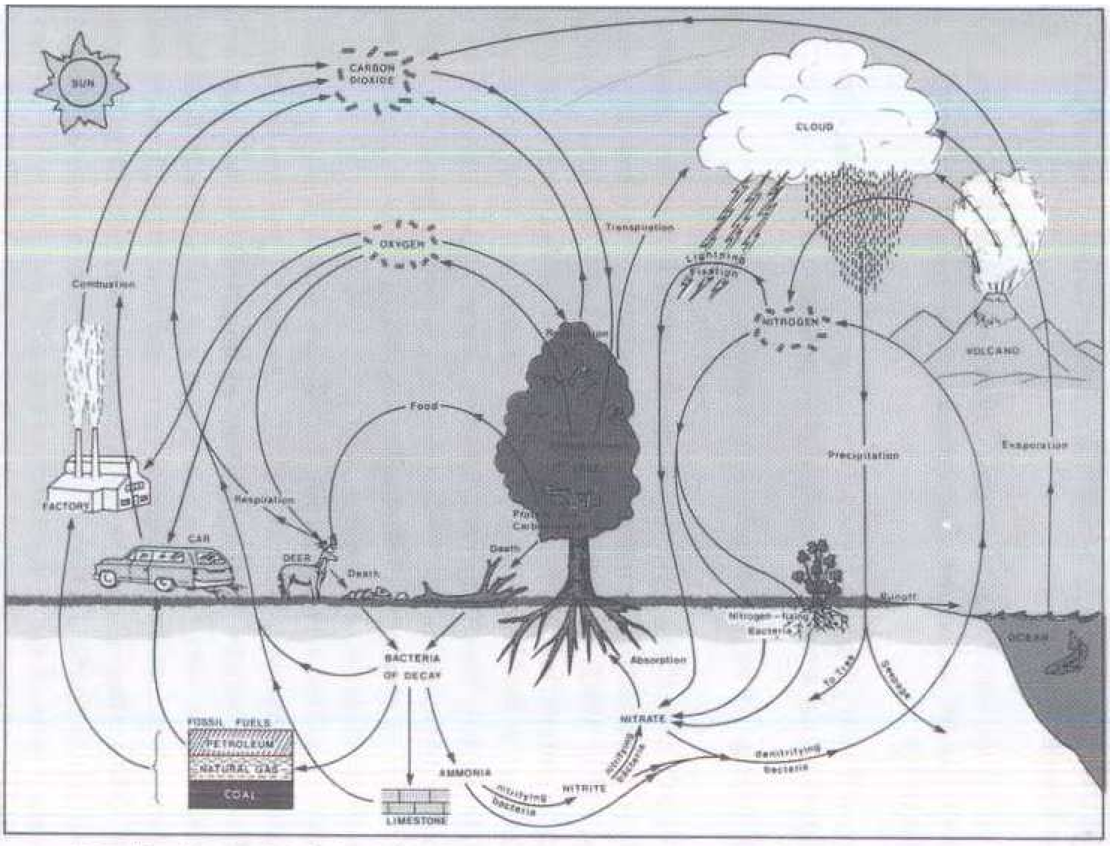


Figure 2.25: Causal diagram about the effects of space travel on the human body used by McCrudden et al. [MSLP07]. It was shown to the participants either with an accompanying text or without text.

Another study was conducted by Branch and Riordan [BR00]. They examined the effects on comprehension using flow diagrams or texts with or without study questions, and different time intervals to study the materials. The participants either had to inspect a flow diagram, which can be seen in Figure 2.26, or read a text containing the same information, and some of them got additional study questions afterwards. These were meant to focus the attention on the most important parts. Students were allowed to spend either 10 or 20 minutes on answering the study questions. According to the results of the post-test, the group which only saw the diagram, performed on average better than the text-only group for a learning time of 10 minutes. However, for a learning time of 20 minutes, it was the opposite.

To summarize, two of the discussed studies [Chi19, MSLP07] mostly found no significant differences between comprehension or recall when using visualizations or texts. The other studies [PR19, BR00] had conflictive results that favored visualizations in some cases and texts in others. The visualizations used in the studies were infographics, causal diagrams, and flow diagrams. Two studies used infographics [Chi19, PR19], and it needs to be mentioned that infographics contain graphics and also explanatory text. Therefore, they might not be suited for directly comparing text to information visualizations or graphs in terms of comprehension or processing speed.



From "Teaching Verbal Chains Using Flow Diagrams and Texts" by W.G. Holliday, 1976, AV Communication Review, 24, P. 69. Copyright 1976 by Author. Adapted by permission.

Figure 2.26: Flow diagram used by Branch and Riordan [BR00]. The participants either saw this diagram or a text with the same information.

2.7 Conclusions from the State of the Art

Due to the ever-growing importance of visualizations, it is crucial to find out how fast visual representations can be processed and how comprehensible they are, especially in comparison to text. While many frequently-used statements about the positive effects of visualizations can be found on the internet, many of them are exaggerated or lack scientific sources for their claims.

Figure 2.27 shows a visual summary of the most important conclusions. Looking at the results of the meta-analyses and studies discussed in this section, it can be said that if there is only a text or a visualization, both seem to be equally well comprehensible. When visualizations are added to texts, they increase the comprehension in most cases. However, it has to be stated that most studies only found very small differences between comprehensibility of texts and visualizations. Some studies even showed that the groups working with texts performed better. It was concluded that people often judge visualiza-

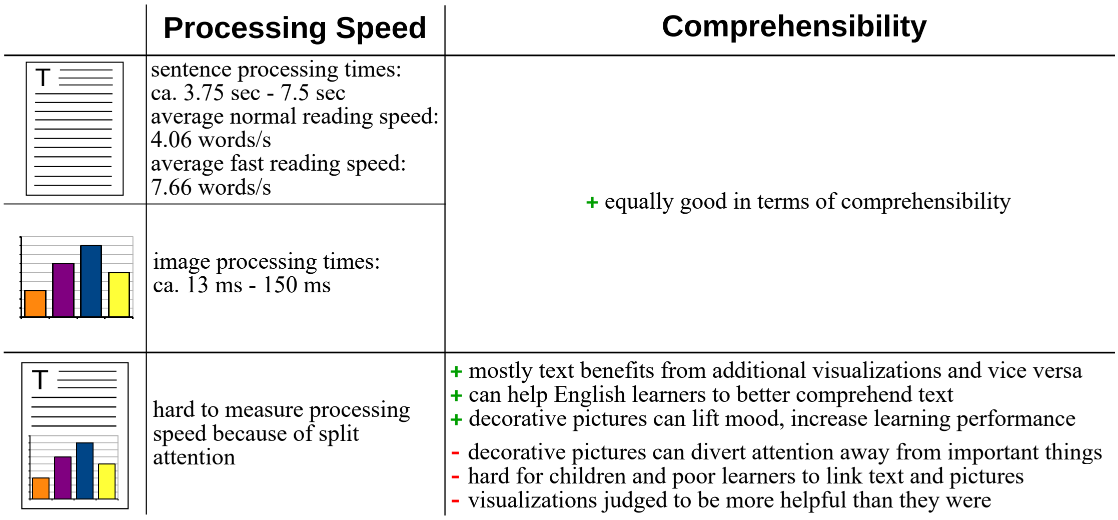


Figure 2.27: Visual summary of the most important conclusions in terms of processing speed and comprehensibility. From top to bottom the conclusions are listed for text-only, visualization-only, and text with additional visualizations.

tions to be more helpful than they actually are. Also, especially for young children, poor learners, and when using decorative visualizations, some negative effects were found.

In terms of processing times, photos seem to be processed faster than texts, but rather 6 to 600 times faster, not 60,000 times. It has to be stated that there is a difference between recognizing a given target in a photo versus being able to interpret, for example, a diagram. So the processing speeds for other visualization types might be very different.

Concerning processing speed, no comparative studies were found at all. In terms of comprehensibility, only four studies directly compared texts to visualizations without text, and they had conflictive results. Also, while the influence of the content was not discussed in the studies, it would be interesting to know which role the content plays, because some information might be easier to be communicated through visualizations or texts. It is also very likely that the type of visualization is important for comprehensibility and processing speed. However, most of the discussed studies only used one type of visualization or compared two different types. Results for more and different kinds of visualizations would be interesting too. Therefore, a study would be required which directly compares text to visualizations in terms of processing times and comprehension, while also analyzing the influence of the type of content and visualization type.

CHAPTER 3

Preparations and Pre-Study

The aim of this thesis was to conduct an exploratory study for the direct comparison of visualizations to texts, both in terms of processing speed and comprehensibility. To achieve this, we first wanted to make sure that the visualizations and texts we use are appropriate, will fit the participants' expectations, and, most importantly, will be comparable. After discussions with other researchers experienced in conducting user studies, we, therefore, decided to set-up a pre-study on how participants would use words and text to describe visualizations. In this chapter we first describe how we selected visualizations and texts for our user study. We further analyze the results we achieved in a pre-study, where we asked participants to use words to describe visual content.

3.1 Choice of Materials

In our exploratory study about processing times and comprehensibility of visualizations compared to text we needed both parts - visualizations and texts representing the same type of information. Similar to the MASSVIS dataset [mas15], we needed to create a dataset for the study where a descriptive text exists for every given visualization. In this section we describe our decisions for selecting visualizations and texts for the study.

3.1.1 Visualizations

An important decision we had to make at this point was the choice of visualization types to be used in our study. While there are also interactive visualizations, where users can filter data, zoom in, or choose certain parts of the visualization to explore them further, we decided to only concentrate on the topic of static visualizations. We disregarded to use scientific visualizations (e.g., volume visualizations), as they require a lot of domain knowledge, and decided to rather concentrate on information visualizations - charts and plots. We also decided to exclude infographics from our study. Infographics were already

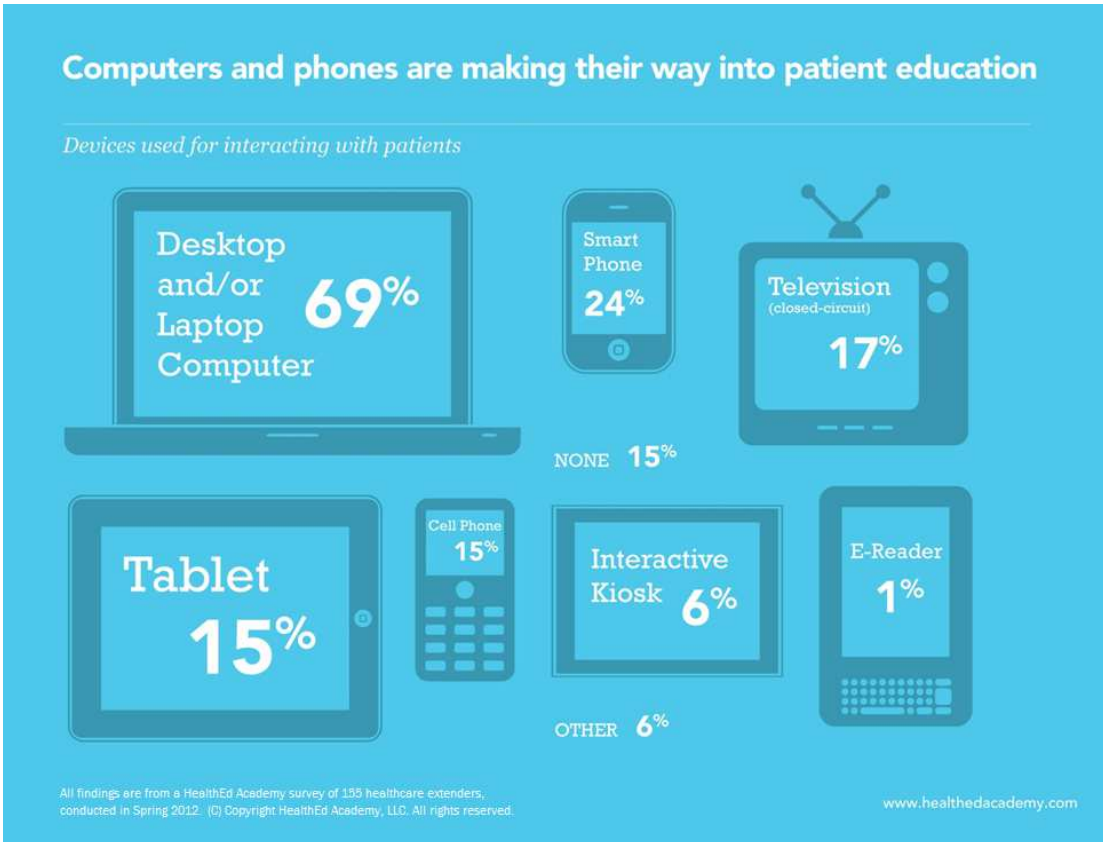


Figure 3.1: Example of an infographic, taken from the MASSVIS Dataset [mas15]. It can be seen that the infographic contains a lot of text, while the visualization parts are mostly decorative and do not add further information.

used in the studies conducted by Chibana [Chi19] and Petrova and Riekhakaynen [PR19] where they compared the comprehensibility of texts and visualizations (see also Section 2.6.4). In contrast to charts, infographics contain both textual and graphical parts. The latter means that the attention has to be split between text and graphic, and it would be impossible to tell the impact of the text or the visuals apart. Eye-tracking might solve this problem, but was not considered in this thesis. Concerning processing speed, the eye-tracking study by Petrova and Riekhakaynen [PR19] revealed that text areas of infographics had more fixations, while some of the graphical areas had no fixations at all. An example of an infographic with a lot of text parts and only decorative visuals can be seen in Figure 3.1.

Guo et al [GZWM20] mentioned that the graphics used for many studies so far were created by experts, and it is, therefore, questionable, if the positive effects that were found also apply to visualizations in general. To increase variability, we took visualizations from different online sources instead of specifically creating visualizations for our study or taking them, for example, from school textbooks or technical books.

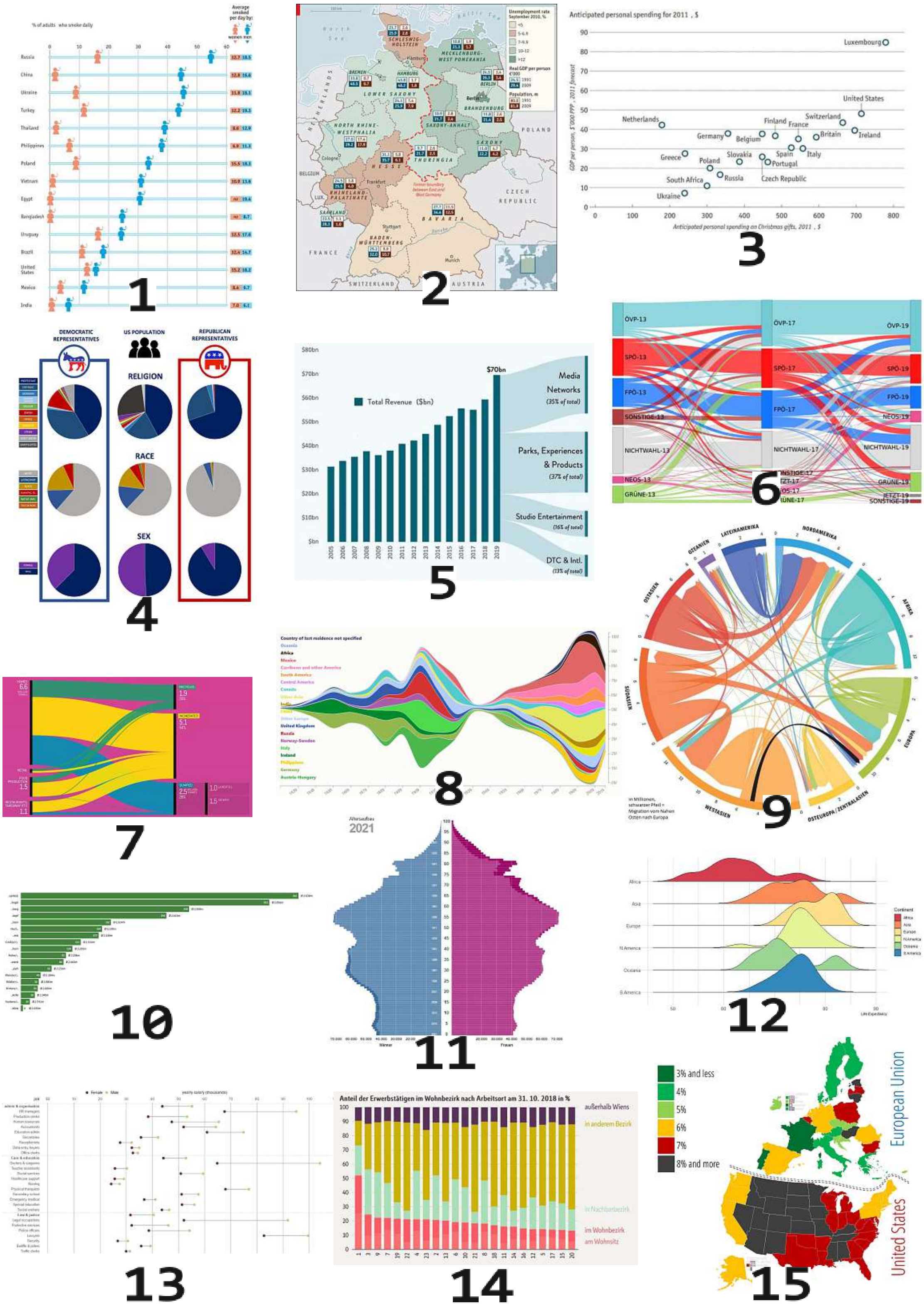


Figure 3.2: The 15 visualizations we used for our study. Visualizations 1, 2, and 3 were taken from the MASSVIS dataset [mas15]. Larger images can be found in Appendix A. Table 3.1 shows the topics and chart types of these 15 visualizations.

We ended up with 12 carefully selected visualizations with different chart types and created with different topics in mind. We also included three visualizations from the MASSVIS dataset [mas15], which increased our final study dataset to 15 visualizations. Thumbnails of the chosen visualizations can be seen in Figure 3.2. Larger images of the visualizations can be found in Appendix A. Table 3.1 shows the topics and chart types of these visualizations.

Nr.	Topic	Visualization Type
1	Cigarette use 2010 [V1]	Comparative dot plot with decorations and additional statistics
2	Unemployment rate Germany 2010 [V2]	Choropleth map with additional statistics
3	Money spent on Christmas gifts 2011 [V3]	Scatter plot
4	Demographics of the U.S. Congress 2020 [V4]	Pie charts
5	Revenue streams of Disney [V5]	Vertical bar chart with Sankey diagram
6	Elections in Austria [V6]	Sankey diagram
7	Food waste in the UK [V7]	Sankey Diagram
8	Immigration to the U.S. 1820 - 2015 [V8]	Stacked area chart/violin plot mixture
9	Global migration 2010 - 2015 [V9]	Chord diagram
10	Austria's mountains [V10]	Horizontal bar chart
11	Austria's population [V11]	Population pyramid
12	Life expectancy per continent 2016 [V12]	Ridgeline plot
13	Gender pay gap in the U.S. [V13]	Comparative dot plot
14	Places of residence and work in Vienna [V14]	Stacked vertical bar chart
15	Yearly inflation 2021 [V15]	Choropleth map

Table 3.1: The topics and visualization types of the 15 visualizations we used for our pre-study, and also later on for the exploratory study. The numbered visualizations can be seen in Figure 3.2 and larger versions in Appendix A.

3.1.2 Texts

One reason why we only chose 15 visualizations was that we needed descriptive text for every single visualization. Every text must contain the same information as the visualization, otherwise they are not comparable. Existing studies show that the formatting, and the length of the text play important roles for its comprehensibility and processing speed.

We compared the lengths of the texts in the studies described in Sections 2.5 and 2.6. Here we realized that the lengths varied greatly, spanning from short paragraphs with less than 100 words to long reports with more than 1,300 words. Only five out of 18 studies used texts that were longer than 500 words, while a big majority of 13 studies used texts with up to 500 words. Of these 13 studies, seven used texts with up to 250 words, and six studies used texts between 251 and 500 words. Two of the shorter texts were only text questions for math tests. As a consequence, we decided that text lengths of 100 to 500 words to be appropriate for our study.

Besides the length we also needed to make sure that texts look approximately the same in all cases. We, therefore, decided not to highlight specific parts of the text, for example, by writing text bold, underlined, in color, or in italics. We only used headlines which had a different format than the rest of the text. This decision was in accordance with most of the studies described in Sections 2.5 and 2.6. Only three of the studies used bold font for highlighting certain words in the texts.

After deciding on the text length and format, one major task to solve was to create texts that contain the same information as the 15 chosen visualizations. That is very similar to the research question when creating the MASSVIS dataset [BBK+15]. For three visualizations we could partly rely on text from the MASSVIS dataset, but for the other 12 we had to create our own texts. When creating such texts, it is very important that they include all the information that is contained in the chart, but without describing every single data point.

Interestingly, there is not much related work available on how to extract text from a visual representation. What can be found are papers and guidelines for alternative image texts, for example, for vision impaired people, see for example the W3C Web Content Accessibility Guidelines (WCAG) on text alternatives [W3C]. Another example are the DIAGRAM Center Image Description Guidelines [DIA]. They provide accessibility guidelines as well as examples for alternative texts for different types of images, including different types of diagrams, graphs, or maps. The clear focus of these guidelines is to provide a very complete picture of what can be seen in the image. This means that alternative texts often contain descriptions of certain attributes like axes of a diagram, colors of bars, or the number of bars, lines, etc. For some diagrams it is suggested to write the data points of the diagram as a list or a table to make them accessible. As such, alternative texts are not suitable to be used in our study.

First we thought we could rely on the descriptive texts from the study conducted by Borkin et al. [BBK+15], since the MASSVIS dataset [mas15] is publicly available. Unfortunately, the longest and detailed descriptions we found in there are made for infographics - a type of representation we decided to discard in our study. For other charts the descriptions were only about 13.1 words (on average). We, therefore, only used three visualization/text combinations from the MASSVIS dataset.

For the remaining visualizations, we had to create our own text representations. From the existing studies (see Section 2.3), we know that people mostly look at extrema, as well as the title of a visualization, and other textual parts. We, however, wanted to avoid any bias in creating the texts ourselves. We, therefore, conducted a pre-study in which we recruited participants and asked them to describe in a textual way (in their own words) what they see in the selected visualizations.

3.2 Pre-Study

Our pre-study was only on how to extract text out of a given visualization. We followed a very similar approach as Borkin et al. [BBK+15]. In this section we describe how we conducted the study and the collected results.

3.2.1 Participants, Materials and Method

The pre-study was conducted as an online user study. Each participant started the study at a page with instructions and afterwards saw four random visualizations out of the 15 that have been chosen (see Section 3.1.1). In a text field, the people had to describe the provided visualization in their own words with at least five sentences. It was possible to describe the visualization either in English or in German. The participants also had to choose on a 4-point Likert scale how confident they were in their description. The possible answers were 'very unconfident (1)', 'slightly unconfident (2)', 'confident (3)', or 'very confident (4)'. Figure 3.3 shows the user interface of the study.

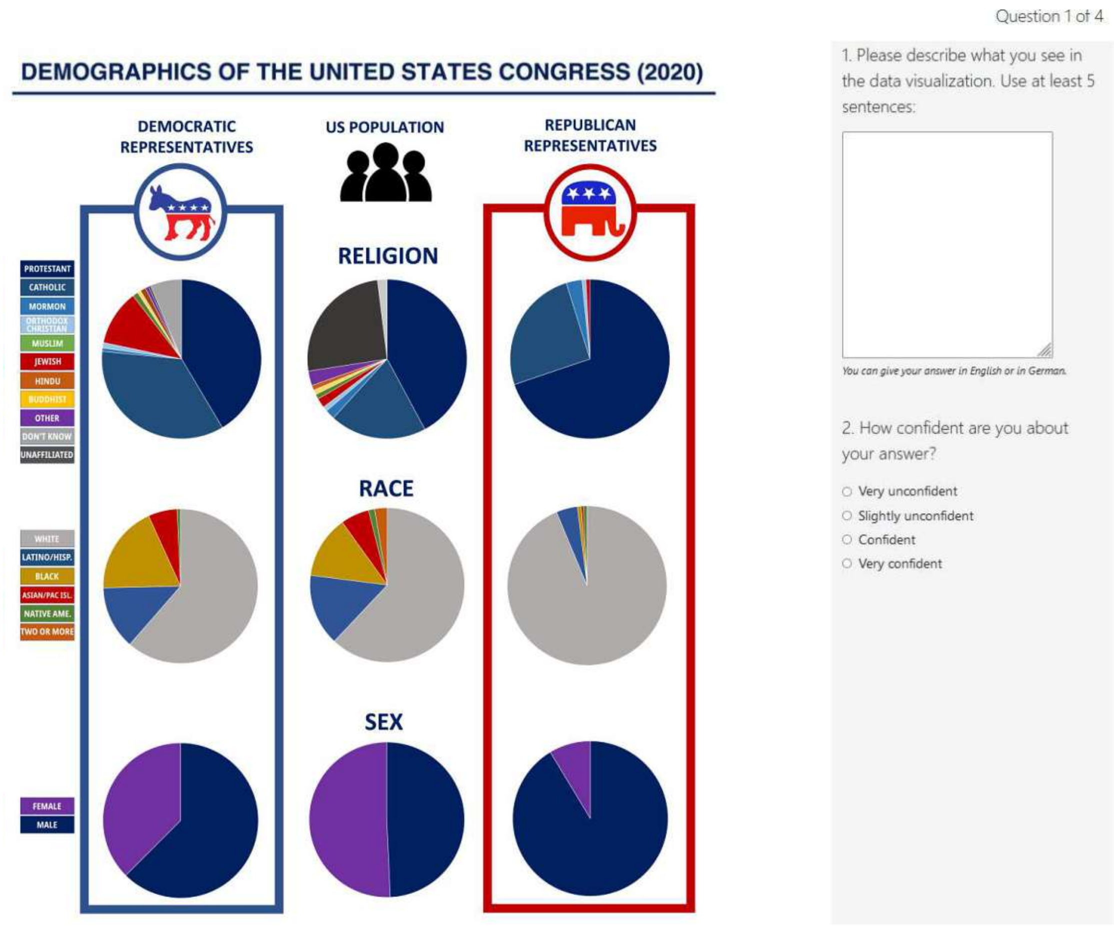


Figure 3.3: The user interface of the online pre-study with one of the visualizations we chose for the study. The randomly chosen visualization was shown on the left side, and participants were asked to enter a descriptive text and their confidence in their description on the right side.

3.2.2 Results

We received five answers for our pre-study. We assume that writing text is quite some effort, and that many people dropped out before completing the exercise. Since we did not use incentives, we did not manage to reach a higher number of participants. Due to this small number of participants we did not receive descriptions for all visualizations, whereas for some visualizations we got more than one description (because of the random selection). In total, we received 20 descriptions for 12 visualizations. Table 3.2 shows how many descriptions we collected per visualization, how long they were, and the measures of confidences for the descriptions. Numbers in braces show the averages in case there was more than one description for a visualization. We also received answers for three additional visualizations we had to later take out of the sample, because we decided to only use 15 visualizations for our study. The three outtakes either contained too much text or other problems, like misleading axis labels that could have confused the participants during the study.

Vis. Nr.	# of Descriptions	Words (Average)	Confidences (Average)
4	3	105, 219, 87 (137)	2, 3, 4 (3)
5	2	98, 167 (132.5)	3, 3 (3)
6	1	66	2
7	2	45, 76 (60.5)	4, 4 (4)
8	1	109	4
10	2	82, 99 (90.5)	4, 2 (3)
11	1	103	4
14	1	211	2
15	1	37	3
16 (outtake)	1	94	4
17 (outtake)	2	76, 56 (66)	3, 2 (2.5)
18 (outtake)	3	201, 52, 47 (100)	2, 3, 3 (2.7)

Table 3.2: Number of descriptions, words and measures of confidences we collected for the visualizations during our pre-study. For visualization number 4, for example, three descriptions were written by different participants with an average description length of 137 words, and an average confidence rating of 3.

The longest description was written for visualization number 4, the visualization about the U.S. Congress. The shortest descriptions were created for visualizations 15 (Yearly inflation) and 7 (Food waste). The confidence measures among the test persons spanned from an average of 2 to 4, where a higher value denoted a higher confidence in their description. No test person assigned the lowest confidence value to any of their descriptions. The distribution of the other three confidence values was quite uniform. 7 descriptions were assigned the highest confidence value of 4, another 7 were assigned a value of 3, and the remaining 6 were rated with a 2. This lead to an average confidence of 3.05 for all of the 20 descriptions. The visualizations where participants assigned on average the lowest confidence values to their descriptions were 6 (Elections in Austria) and 14 (Places of residence and work in Vienna). The highest average confidence values for descriptions were assigned for visualizations number 7 (Food waste), 8 (Immigration to the U.S.), 11 (Austria's population), and 16 (outtake).

The contents of the descriptions were analyzed manually. We noticed different writing styles, which are also reflected in the length of the descriptions. However, there were common terms and features that were detected by all participants. Table 3.3 shows which elements were mentioned in the descriptions per visualization. The numbers in braces denote by how many test persons an element was mentioned, in case it was mentioned more than once. The elements are ordered by the number of occurrences. For all but one visualization either the general topic of the visualization or the title were mentioned. It has to be stated, that not all of the visualizations had titles.

Vis. Nr.	**Elements mentioned in the Descriptions (# of Descriptions in which it was mentioned, if more than 1)**
4	Comparisons between Values (x3), Topic (x2), Extrema (x2), Visualization Type (x2), Axis Descriptions (x2), Party Logos, Legend
5	Topic (x2), Extrema (x2), all Values of Sankey Part (x2), Title, Visualization Type, First Value, Comparisons between Values
6	Topic, Axis Descriptions, Extrema
7	Extrema (x2), Topic, Comparisons between Values, Visualization Type, Axis Descriptions, all of the Values on the Right Side of the Chart
8	Visualization Type, Topic, Extrema, Shape, Comparisons between Values
10	Topic (x2), Axis Descriptions (x2), Relationship of Axes, Extrema
11	Topic, Axis Descriptions, Extrema
14	Visualization Type, Title, Axis Descriptions, Extrema, all Stacked Values from first Bar, Comparisons between Values
15	Comparisons between Values

Table 3.3: The elements that were mentioned in the descriptions of the different visualizations ordered by number of occurrences. For example, three participants included comparisons between certain values in their description of visualization number 4. Elements that were mentioned in two of the descriptions of visualization number 4 were the topic, extrema, visualization type, and axis descriptions. Party logos and the legend were mentioned in one of the descriptions of visualization number 4.

Figure 3.4 shows the total number of occurrences of the different elements that were mentioned in the descriptions by the test persons. The elements that occurred most often were extrema, which were mentioned in 17 of the 20 descriptions, followed by the general topic of the visualization, and value comparisons, which were both mentioned in 13 of the 20 descriptions. The fact that people are mostly concentrating on extrema and labels confirms the results of the study by Borkin et al. [BBK+15]. An interesting finding was that people liked to describe the visualizations by using comparisons between different

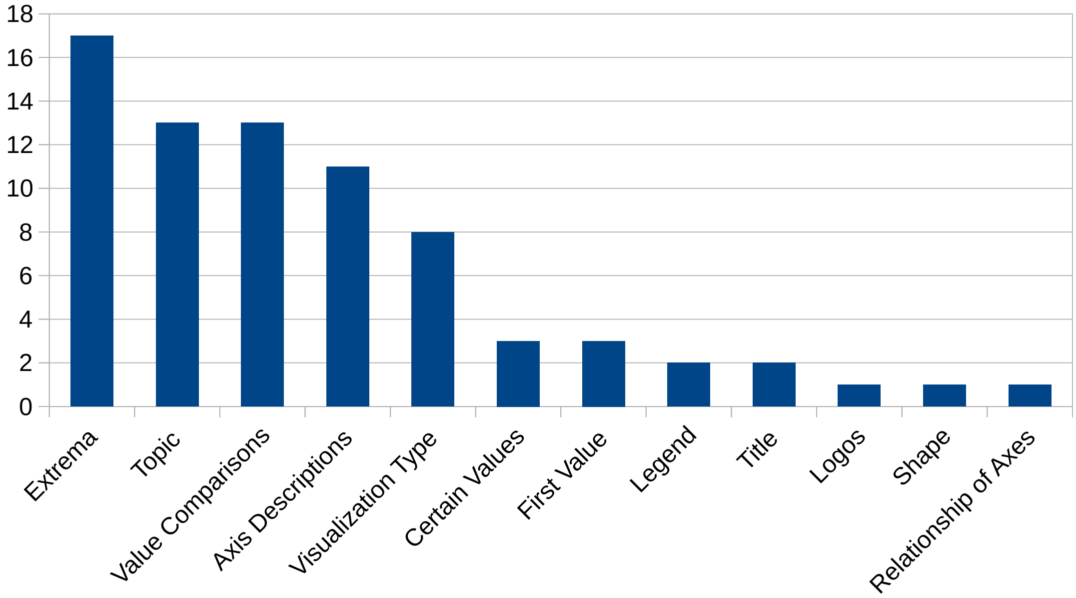


Figure 3.4: Number of occurrences for the elements mentioned in the descriptions. Extrema were the most mentioned elements. They occurred in 17 of 20 descriptions, followed by topic of the visualization and comparisons between values (both in 13 of 20 descriptions). Logos, shape of the visualization, and relationship of axes were only mentioned in one description each.

values. For example, if there was a bar chart where the bars denoted the values for different years, people did not just write down the numbers for certain years, but wrote that the numbers increased or decreased in comparison to other years. Axis descriptions and the visualization types also got mentioned quite often. In some cases, the value of the first element in a visualization was mentioned, even if it was not an extremum, e.g., the value of the first bar in a bar chart. In some cases certain values, for example, all values on the right side of the Sankey Diagram about food waste (visualization number 7), were mentioned. Also, both descriptions about visualization number 5 (which is a bar chart where the bar on the right is partitioned into a Sankey Diagram) contain all values of the Sankey Diagram part. Elements that were mentioned less often were legends, logos, the shape of a visualization, and the relationship of values on different axes.

3.2.3 Conclusions

The pre-study confirmed the conclusions of the study by Borkin et al. [BBK$^+$15] that stated that extrema get much attention from the viewers. We found that people tend to describe visualizations by using comparisons and relationships, rather than just listing the values.

CHAPTER 4

Study

The main aim of this thesis was to conduct an exploratory study on the processing speed and comprehensibility of information visualizations (charts, plots) versus texts. We are aware that time and accuracy alone are probably not possible to capture all parameters of an interactive visualization [NWC+21]. However, our visualizations are not interactive, so we concentrated on time and accuracy only. In this chapter we discuss the baseline of the study, describe how we conducted the study, and discuss the outcomes.

4.1 Research Questions

Our exploratory study was based on three research questions which we wanted to explore further:

Research Question 1: Do humans process visualizations or texts faster, and how much faster?
As discussed in Section 1.1, a widespread claim states that visuals are processed 60,000 times faster than text. Studies about processing speed suggest that the difference is likely much lower and rather between 6 and 600 times. However, none of the studies we found directly compared the processing speed of visualizations to that of texts, so it is still unclear which medium is faster and how much faster.

Research Question 2: Do humans comprehend information from visualizations or texts better, and how much better in terms of accuracy?
The results from the studies discussed in Section 2.6 suggest that adding meaningful visualizations to texts increases comprehensibility and has some other positive effects, too. However, we did not find many studies directly comparing comprehensibility of visualizations to that of texts. Studies that did have inconclusive or conflictive results suggesting that visualizations and texts are comprehended equally well.

Research Question 3: Does the content play a role for processing speed or comprehensibility?
The studies discussed in Sections 2.5 and 2.6 did not really pay attention to the question if the content of a visualization or a text plays a role for the processing speed and/or comprehensibility. Since we consider this a relevant and interesting question to be explored, we decided to include this research question in our exploratory study.

The aim of our exploratory study was to get initial results for the proposed three research questions. Based on the results, we would be able to define hypotheses and directions for further work [Ste01].

4.2 Materials

We used the same 15 visualizations that were also used for the pre-study. The visualizations can be seen in Figure 3.2. For every visualization we needed to define an equally descriptive textual part. We used texts from the pre-study and also formulated our own textual representations. Finally, we ended up with 15 visualization/text pairs. They can be found in Appendix A. Please note that from now on such pairs are referred to as 'topics'. A list of all topics can be seen in Table 3.1. We made sure to remove all the texts from the visualizations that were unnecessary for their comprehensibility. This included additional informations about the topic in the visualization, logos, sources, as well as the titles. The titles were removed because not every visualization had a title. During the study, visualizations and texts would be presented with uniform titles, so that each visualization and the corresponding text would have the same title. Also, by presenting them like this, the titles for all the texts and visualizations were written in the same font, as well as the same font size.

As mentioned in Chapter 3, we conducted a pre-study to create texts out of the given 15 visualizations. Since in the pre-study we could not collect enough material, we had to manually create the texts ourselves. We wrote corresponding texts in German and translated them to English, according to the findings from the pre-study. We made sure to include all the information of a visualization from the descriptions that were written by the participants in the pre-study. Our texts also included extrema and the main topic of the visualization, as well as comparisons between different data values of the visualization. All of the texts can be found in Appendix A. Figure 4.1 shows the number of words for each of the texts in English and German. All of our texts were between 194 and 411 words long. The average text lengths were 314.53 words for the English texts and 300.53 for the German ones. The topic with the shortest English and German texts on average was topic number 5 (Revenue streams of Disney), and the topic with the longest texts on average was topic number 9 (Global migration 2010 - 2015).

We did not aim at creating texts with uniform lengths for all given visualizations. The different text lengths mainly reflect the different complexities of the corresponding visualizations and the represented information. The main goal of the study was to compare visualizations to texts, so we assumed that more complex visualizations (which

Figure 4.1: The number of words of our English and German texts per topic. A list of all topics can be seen in Table 3.1. The shortest texts were written for topic number 5 and the longest ones for topic number 9.

need more text to describe) will also take more time to be understood. Since we did not compare the different texts in the study, the different lengths did not play a major role.

An example for a visualization and text pair for topic number 7 (Food waste in the UK) can be seen in Figures 4.2 and 4.3. The information displayed informs the viewer/reader about food waste in the UK.

For testing comprehensibility, we decided to ask the participants three questions after showing them either a visualization or a text. So, beforehand, we had to formulate three questions for every visualization/text pair we had. The questions were about finding extrema, retrieving certain values, and comparing different values. Questions were available in German and English. The questions can be found in Appendix A. Regarding topic number 7, which is shown in Figure 4.3, the following questions have been phrased:

1) From which source is the second most food recycled?

2) How many tonnes of food end up in the sewer?

3) What is the percentage difference between the amount of food that is incinerated and food that is recycled?

All in all more than 9 million tonnes of food get discarded every year in the UK. More than half of the discarded food comes from homes. Thus, homes are by far the largest cause for food waste with 6.6 million tonnes per year. A smaller portion of food waste is generated by food production with 1.5 million tonnes per year, followed by restaurants and takeaways with 1.1 million tonnes per year. Only a very small and neglectable percentage of food waste stems from retail. 54% or 5.1 tonnes of the waste food from all causes are incinerated. Some of this is used to generate electricity or is burnt to get fuel, but some is also released into the atmosphere. More than half of the incinerated food waste stems from private homes, followed by food production and restaurants and takeaways. A small percentage stems from retail. 26% or 2.5 million tonnes of food get dumped, of which 1 million tonnes land in landfill and the bigger portion or 1.5 million tonnes is dumped into the sewer systems. The problem with this is that the rotting food in landfills generates methane, which is a greenhouse gas. Also, dumping food into the sewer system can lead to problems like blockages and damaging of the pipes, which can cause high costs of repair. More than 2 million tonnes of the food that gets dumped stems from private homes, a small percentage from restaurants and takeaways and a very small percentage from food production. Only 20% or 1.9 million tonnes of food get recycled. This food is then redistributed to food banks and vulnerable people. Most of it comes from private homes that only recycle roughly 15% of their waste, followed by food production, which recycles nearly one third of its food waste, and retail, which recycles half of its waste. A small portion is from restaurants and takeaways.

Figure 4.2: The text for topic number 7 about food waste.

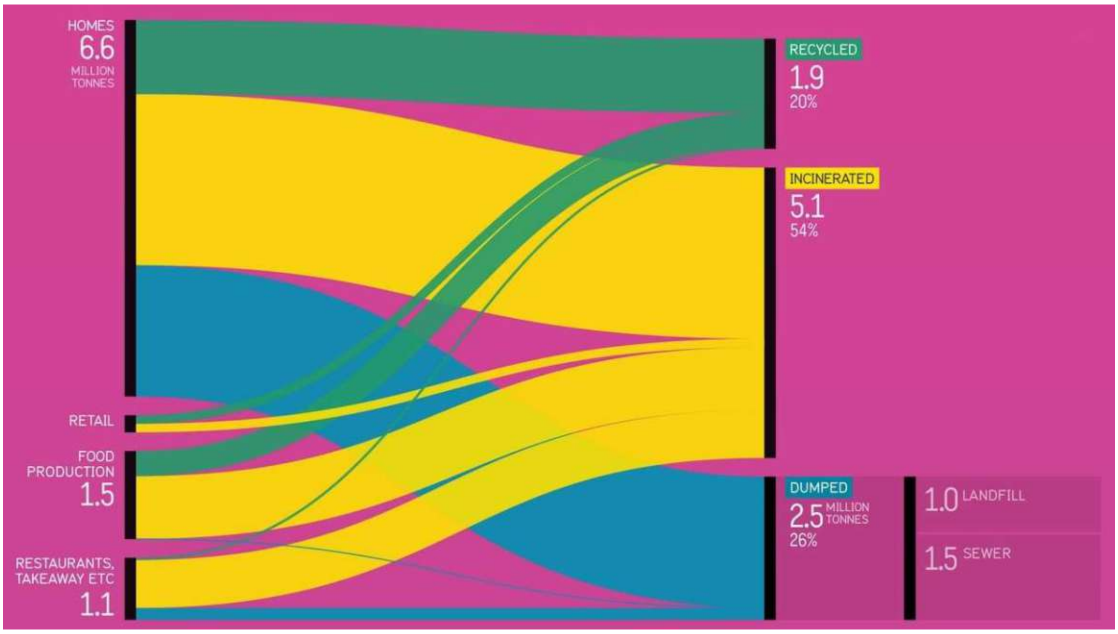


Figure 4.3: The visualization for topic number 7. It displays how much food is wasted in the UK, from which sources it comes (left side of the diagram), and what happens to the wasted food (right side of the diagram).

4.3 Participants

All participants in our study were adults that had not participated in the pre-study. The latter was important, because if a participant had already seen one of the visualizations, the results would have been distorted, as they maybe would have needed less time to comprehend the visualization. We had 18 participants, of which 11 were female, and seven were male. The youngest participants were 22 years old and the oldest 60 years with a total mean of 35.3 years. The average age of the male participants was 31.14 years, and 38 years for the female participants. Participants had to be fluent in either English or German in order to be able to conduct the study. Four of the participants chose to conduct the study in English, while 14 conducted it in German. Only one participant reported that he/she sometimes has problems with distinguishing very similar colors. Post-checks of the results and comparison with the other participants did not suggest that he/she was influenced by this during the study. No other visual problems of participants that could possibly influence the results of the study were known.

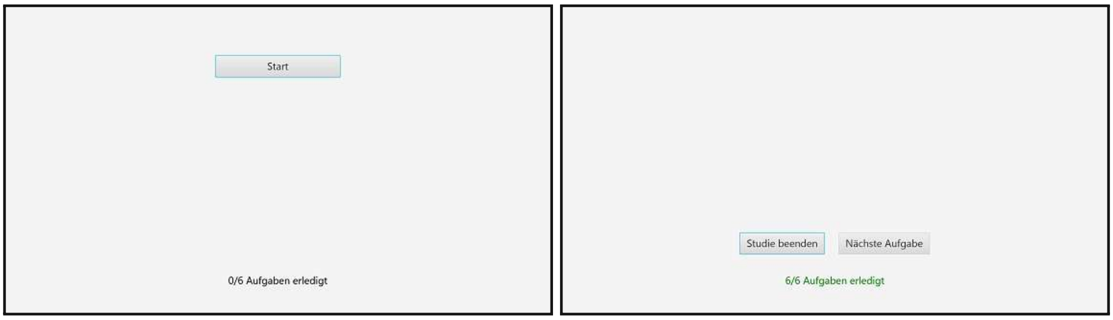
4.4 Methodology

The study was conducted in a supervised way. We decided against an online setting to be in control of how the visualizations are presented. We implemented a standalone study software which is described in Section 4.4.1. The study procedure is described in Section 4.4.2. Throughout the study participants had to solve certain tasks. The term 'task' now refers to seeing a visualization or text (from a certain topic), understanding it, and answering three questions about it. The order of the tasks was chosen randomly. However, we made sure that if a visualization was shown, the participants would not see the corresponding text during the study - and vice-versa. Participants had to solve at least six tasks. After the sixth task, participants could freely decide to go on, or to end the study.

4.4.1 User Study Framework

We implemented our own software for conducting the user study. When started, the software ran in full screen mode without showing any menu bars, so the participants could not accidentally close the study. First, the participants were presented with a screen where they had to fill out their age and gender and choose the language for the texts and questions. The background color and the whole design of the program were intentionally kept in inconspicuous colors, so nothing would divert the participants from concentrating on the visualizations, texts, and questions.

After filling out the form and clicking on the button, the participants were presented with the 'start screen' shown in Figure 4.4 (a). After clicking on the 'Start' button, the study started and participants were presented with the first task. At the bottom of the screen there was a counter so that the participants could see how many tasks they already solved, and how many more to go. This screen was also always shown between



(a) (b)

Figure 4.4: (a) shows the screen at the start of the study, as well as between the topics. (b) shows the screen after six topics were solved.

the tasks, so the participants could decide when to begin with the next task. After the sixth task, i.e., when the participants solved the minimum number of tasks, the screen changed, which can be seen in Figure 4.4 (b). Instead of the 'Start' button it showed two buttons so the participants could decide to finish the study, by clicking on the left button, or to solve another task by clicking on the right button.

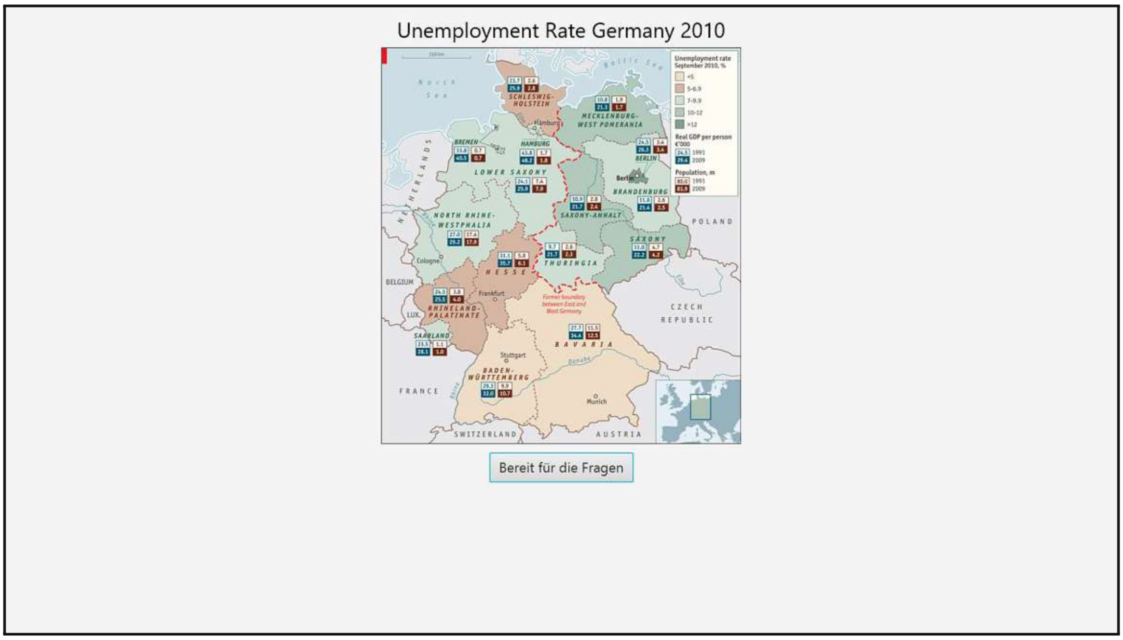


Figure 4.5: An example of how a visualization task was presented to the participants during the study. First, only the visualization and title were shown. By clicking on the button beneath it, the questions were shown (see Figure 4.6).

Figures 4.5 and 4.6 show an example of a task where participants were presented a visualization. First, only the title and visualization were displayed, as can be seen in

Figure 4.5. Below the visualization, there was a button the participants had to click when they felt ready to answer the questions. Figure 4.6 shows the resulting screen with visualization, and questions, both presented at the same time. The order of the three questions per topic was the same for all of the participants. Below each question, there was a text field to fill in the answer (1-5 words). After answering the questions, the participants had to click on the button below the questions to finish the task. This brought them either to the screen in Figure 4.4 (a) in case they did not solve at least six tasks yet, or, to the screen in Figure 4.4 (b) in case they were already beyond the sixth task. The screen for a text task can be seen in Figure 4.7. We disabled the possibility to highlight certain parts of the text, search in the text, or copy parts of the text in order not to favor participants who know certain key combinations (e.g., copy-paste).

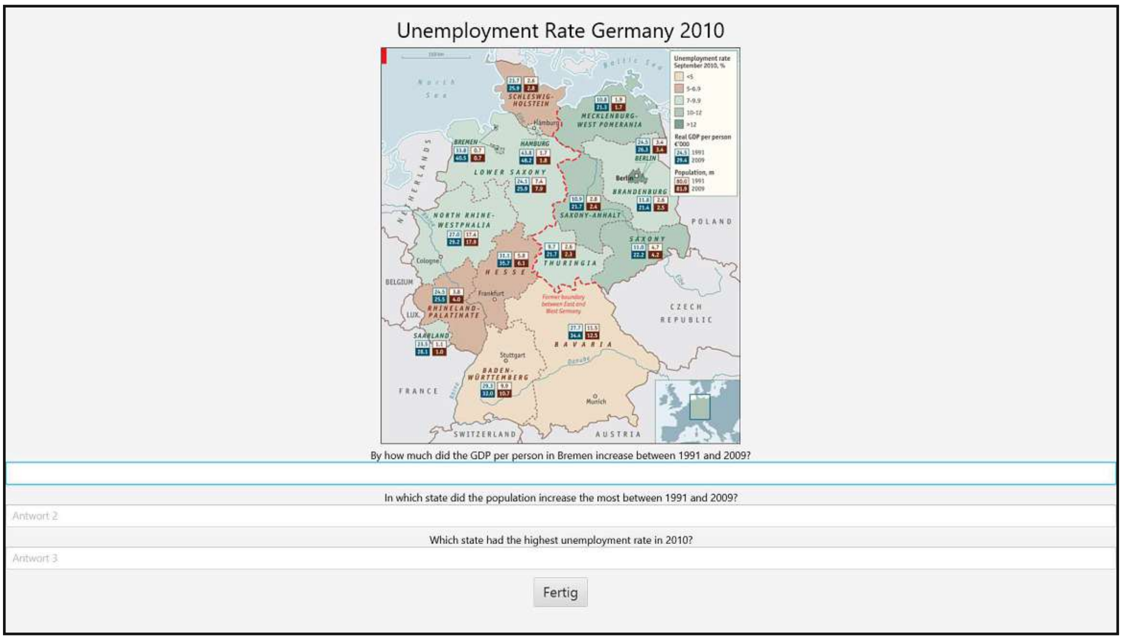


Figure 4.6: This screenshot shows how a visualization task was presented together with the three questions. After the participants had finished answering the questions, they had to click on the button below the questions to finish the task and get back to one of the screens shown in Figure 4.4.

The program took two time measurements for each task. Firstly, it recorded the time between clicking on the ‘Start’ button, and clicking on the button to show the questions. Secondly, the time was measured between clicking on the button to show the questions, and clicking on the button to finish the task. We wanted to differentiate between the time the participants looked at the visualization or read the text, and the time it took them to answer the questions. After clicking on the button to end the study, a JSON file with the data of the participants, their answers, and their timings was created and saved. Since it was an anonymous study, the JSON file also contained a randomly generated ID for each participant so we could distinguish between them.

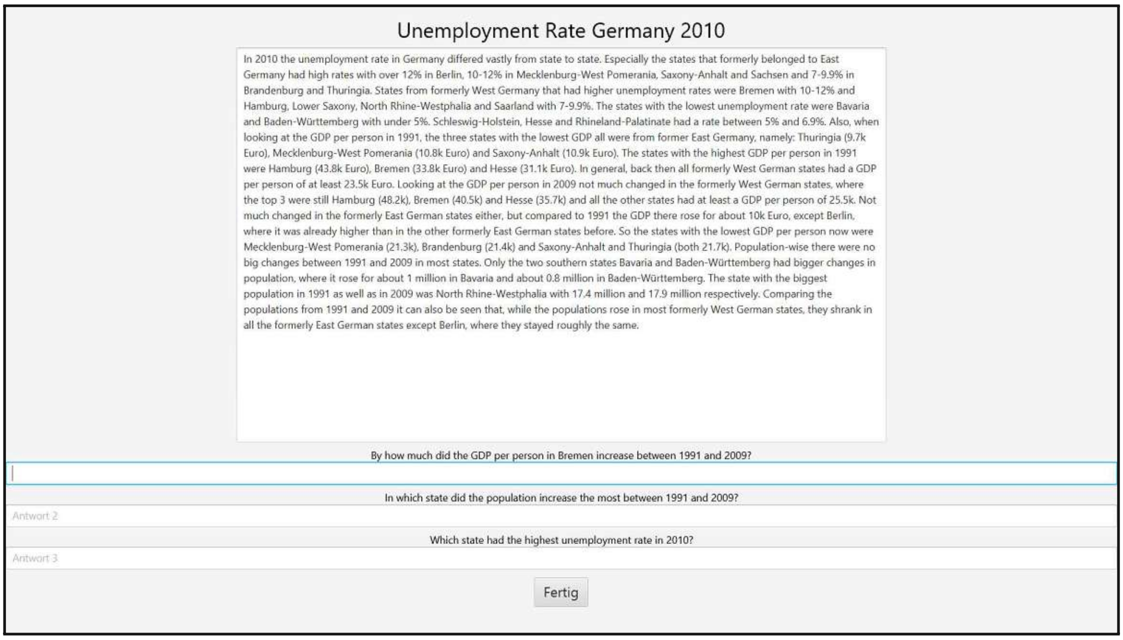


Figure 4.7: An example of how a text task was presented during the study. This screenshot shows the same topic as given in Figure 4.6, but in text form. Title and questions are the same as for the visualization task.

4.4.2 Study Setting

In the beginning we planned to recruit at least ten participants for our exploratory study. In order not to make the same mistake as in the pre-study, where we did not get descriptions for every visualization, we, this time, made a task plan. This task plan guaranteed that, if we had ten participants, every visualization, as well as every corresponding text, was seen by at least two participants. The order of the tasks was chosen randomly, but within the first six tasks, no topic was presented twice. This means, for example, that topic number 5 would not be presented as the first task to two different participants. After the sixth task, the order was also random, but since we did not know if any of the participants would solve more than six tasks, and which participants would choose to do so, we did not make sure anymore that two topics were not at the same position. The task types, i.e., visualization or text, were predetermined for the first six tasks according to the task plan. After the sixth task per participant, they were chosen randomly.

The study was conducted in a controlled environment, in a dedicated office at the premises of the research partner. We invited the participants to the office, so they would have roughly the same conditions when doing the study. The study was conducted by running the program on a PC connected to a 27" monitor. Two participants conducted the study on a notebook with a 23" screen.

Before starting, we explained the purpose of the study. Afterwards, we showed the participants a test study using the same software as the one later used in the study, but it only contained test images and test texts. While explaining to them the functionality of the software, the participants could try out the buttons and get accustomed to the study software. We told them that the time was not measured between the topics, and that they could take a break when seeing the 'start screen', ask questions, or talk to us. However, we also told them that after they clicked on the button to start a topic, they should really focus on the task as time will be measured then. We asked participants to always fill in the answer to the questions in every topic. If not able to answer, we asked them to write either their closest guess, or the reason why they could not answer the question (e.g., values were too similar). Only the supervisor and the participant were in the same room when conducting the study, to avoid any disturbances or interruptions.

After solving the topics, we asked participants some questions about the study where we took notes. These were the questions:

1) What do you think worked better for you, was better understandable/faster? Visualizations or texts?
2) Were there certain visualizations or texts that you found more difficult than others?
3) Were there certain questions you found harder to answer, or which you could not answer? What do you think was the problem?
4) Did you find the visualizations or texts more interesting?
5) Do you think that a certain visualization type was better understandable than others? If yes, which one?
6) Was there a visualization that you particularly liked/that was particularly interesting for you? Which one?
7) Do you usually like to read? Do you tend to read fast or slow?
8) Did you double check the answers before clicking 'Done'?
9) Do you have any other comments, things you want to discuss, or things we can improve about the study?

4.5 Results

In total, our participants solved 193 tasks, so 85 more than they were required to solve. 13 of the 18 participants decided to solve more than the required six tasks. Eight of them even looked at all of the 15 topics. On average, the participants solved 10.7 tasks. From the 193 tasks, 98 were visualization tasks and 95 were text tasks. It can be seen that the distribution between visualization and text tasks was almost even. Table 4.1

Topic	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Seen as visualization task	5	5	10	8	5	8	6	9	6	5	6	7	6	4	8	**98**
Seen as text task	10	7	6	4	6	5	5	5	5	7	6	6	6	9	8	**95**
Total	**15**	**12**	**16**	**12**	**11**	**13**	**11**	**14**	**11**	**12**	**12**	**13**	**12**	**13**	**16**	**193**

Table 4.1: The table shows the 15 topics, and how often they have been seen by the participants as visualization or as text task. Topic number 1 has been mostly seen as text task, while topic number 3, for example, has been mostly seen as visualization task.

shows the distribution of visualization and text tasks among the chosen 15 topics. Topic number 1, for example, has been seen more often as a text task. Topic number 3 has been seen more often as a visualization task. For the other topics, the distribution was quite even. We ensured that every visualization and every text were seen by at least four participants. Every participant saw at least three texts and three visualizations. We also ensured that every visualization and text was seen by at least one male and at least one female participant.

4.5.1 Comprehensibility

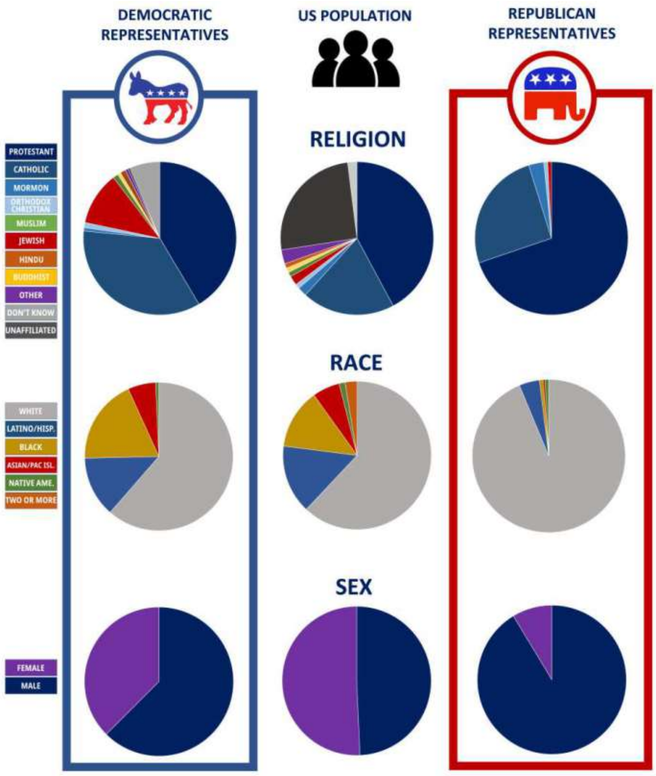


Figure 4.8: An example for how we evaluated the tasks showing the visualization for topic number 4. The question was to read the percentage of the light blue segment in the pie chart in the upper right corner. The correct answer was 25 %.

One major goal of the study was to test the comprehensibility of visualizations vs. texts. The answers to the questions were used to measure comprehensibility and were evaluated manually. We went through all answers and assigned either 0, 0.5, or 1 point per question. 0 points were assigned if the answer was wrong, a very bad estimation of a value was given, or the participants did not know the answer. 0.5 points were assigned if the estimate was slightly wrong, and 1 point was assigned for correct answers. For an example we would like to refer to topic number 4, whose visualization part is displayed in Figure 4.8. One of our questions was to read the percentage displayed by the light blue segment in the pie chart in the upper right corner. The correct answer was 25 %, for which we assigned 1 point to the answer. An estimate between 23 % and 27 % (i.e., up to 2 % more or less than the correct answer) was assigned 0.5 points. An estimate of 29 % was considered wrong and led to 0 points being assigned.

Three points maximally could be achieved per task, since we asked three questions in every task. A participant who solved six tasks could, therefore, get a maximum of 18 points. For further evaluation, we summed up the points a participant achieved for each task and calculated the percentage of points compared to the maximum of available points. For example, if a participant achieved 13.5 points of a total of 18 points for six tasks, we assigned a comprehensibility score of 75 %. Also, we rated this as 75 % of the participant's answers being correct and 25 % being incorrect.

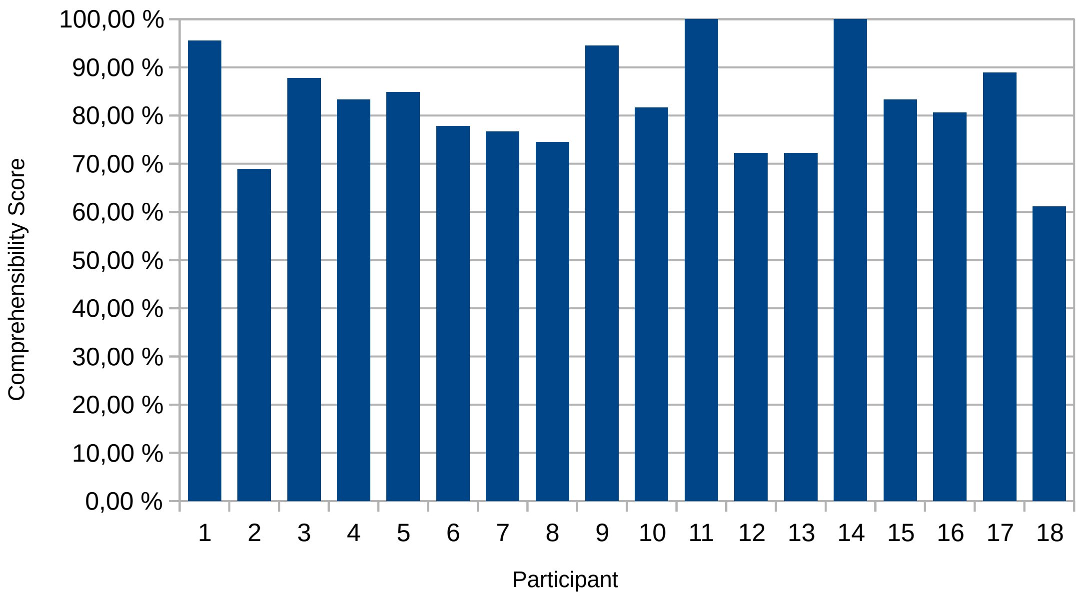


Figure 4.9: The comprehensibility scores for the 18 participants. Participants number 11 and 14 made no mistakes in answering the questions and, therefore, reached a comprehensibility score of 100 %. All scores were between 61.11 % and 100 %.

Figure 4.9 shows the comprehensibility scores for the 18 participants. Participants scored more points when they read the texts (M = 0.84, SD = 0.11) as compared to when they saw visualizations (M = 0.81, SD = 0.13). However, the difference is rather small and also not statistically significant ($p = 0.40$). In total, 18.91 % of questions were answered incorrectly, thereof 20.07 % with visualizations and 17.72 % with texts. Only eight questions were answered correctly by all of the participants.

Two participants, i.e., participants number 11 and 14, made no mistakes and, therefore, had a comprehensibility score of 100 %. Comprehensibility per participant ranged between 61.11 % and 100 %, with an average of 82.43 % in total. Male participants had a comprehensibility score of 87.04 % on average (M = 0.87, SD = 0.08), while female participants had a comprehensibility score of 79.49 % on average (M = 0.79, SD = 0.11). The differences were quite large, but were not statistically significant ($p = 0.14$). The results became more interesting when looking specifically at the text tasks. Here the male participants had a comprehensibility score of 90.37 % (M = 0.9, SD = 0.07), which is quite high. Female participants scored 80.08 % (M = 0.8, SD = 0.11). This difference is also statistically significant ($p = 0.04$) and a very interesting result. It indicates that male participants gave significantly more correct answers when reading texts, than women.

No statistically significant difference could be found between participants being experienced in working with visualizations (comprehensibility score of 82.66 % (M = 0.83, SD = 0.08)) and people working in other fields (comprehensibility score of 82.06 % (M = 0.82, SD = 0.13)). The statistical test led to a p-value of $p = 0.92$.

Again, no significant differences for comprehensibility could be found between participants who solved the tasks in English (M = 0.86, SD = 0.1), or in German (M = 0.81, SD = 0.1). The p-value was $p = 0.53$ in this case. Also age did not make a difference ($p = 0.15$), or if the participants solved less than 10 tasks ($p = 0.16$).

Figure 4.10 shows the percentages of questions per topic that were answered incorrectly with visualizations and texts. We calculated these percentages by summing up the points all the participants achieved per question for each topic and dividing it by the number of solved questions per topic. Every topic consisted of three questions. For example, if one topic was solved by 2 participants and they each answered 1 question incorrectly, they achieved a sum of 4 points for 6 questions. So we calculated a score of 67 % correctly answered questions and, likewise, 33 % incorrectly answered questions for this topic. Topics 4 and 13 seemed to be especially challenging. For topic number 4, participants had comprehensibility problems when they had to solve the visualization task. This was exactly the other way round for topic number 13. Here participants had bad comprehensibility scores when they had to solve the text task. Topic number 9 was the one where we received the most incorrect answers. Another topic with many incorrect answers for the visualization task was topic number 10. For topic number 9, when solving a text task, participants performed nearly equally bad as with the visualization, leading to the second worst comprehensibility score. Most correct answers were given for topics 5, 8, and 12. Concerning visualization tasks, we did not receive any wrong answers for topic number 5, which made it the most comprehensible one with visualizations, followed by topics number 2 and 1. The best comprehensible texts were for topics number 15, 10, and 12.

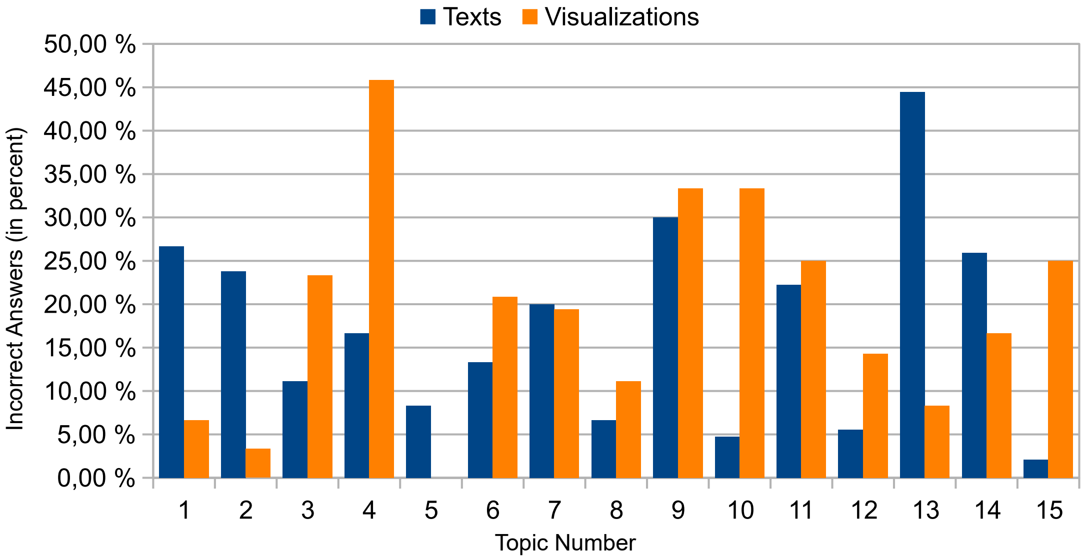


Figure 4.10: The percentages of incorrect answers per topic for text tasks (blue bars) and visualization tasks (orange bars). The topic with the highest percentage of incorrectly answered questions was topic number 9 with 30 % of incorrectly answered questions for texts and 33.33 % for visualizations. Topic number 5 was the topic with the lowest percentage of incorrectly answered questions, since all questions for the visualization were answered correctly, and for texts only 8.33 % of questions were answered incorrectly.

Figure 4.11 shows the average correctness for the first to the tenth task the participants solved. Correctness is shown for visualization tasks, text tasks, and in total. There were some increases and decreases in total, for visualization tasks, and text tasks without showing a general trend. From this we can deduce that the order of the tasks did not have any influence on the correctness of the answers, i.e., the quality of the answers did not rise or drop in the course of the study.

To learn more about incorrect answers, we investigated two of the topics with the most incorrect answers in total and biggest differences between correctly answered questions with texts and visualizations, namely topic number 4 and 13, in more detail. The main problem with topic number 4 was the visualization task, while for number 13 it was the text task. Both texts had similar lengths and were among the five longest texts in the study. The visualization of topic number 4 displayed nine pie charts that had to be compared to answer the questions. The difficulty in comprehensibility might have been that participants had to guess the percentages denoted by the pie segments, because they were not labeled, while they were clearly stated in the text. The visualization of topic number 13 was a comparative dot plot. The text contained many numbers that had to be compared and participants also had to do some calculations. It looks like this task was much easier to solve with the dot plot.

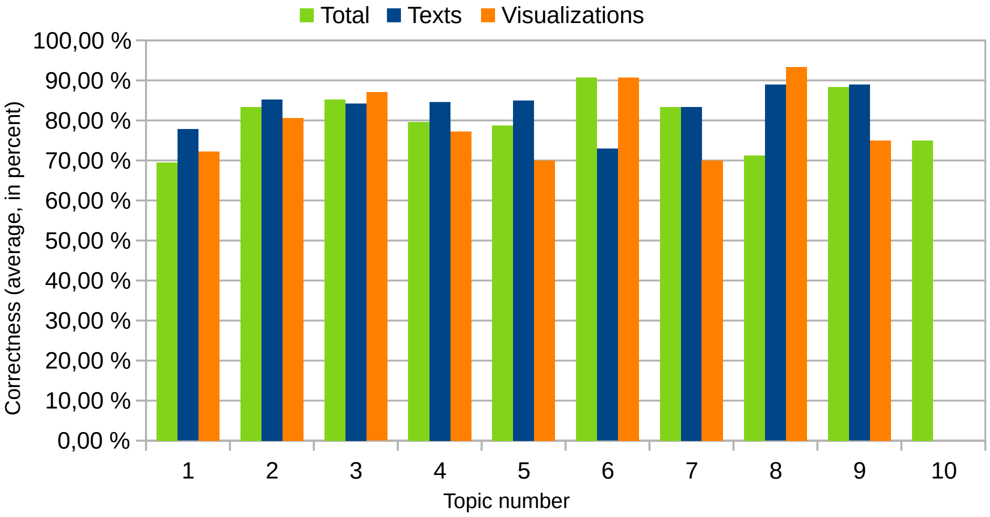


Figure 4.11: Average correctness of the participants for the first to tenth task in total (green), for text tasks (blue), and visualization tasks (orange). Since the maximum number of visualizations or texts a participant saw was nine, there are no blue or orange bars for the tenth topic. The average number of tasks that were solved per participant was 10.7, so there are ten green bars. It can be seen that the order of the tasks did not have any influence on the comprehensibility scores of the participants as there were some increases and decreases in correctness without showing a trend.

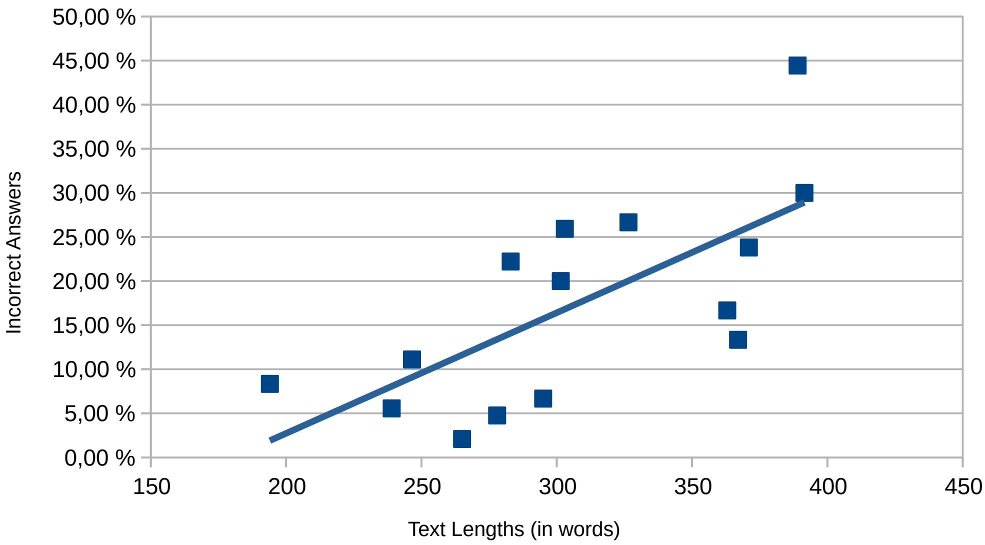


Figure 4.12: A scatter plot with regression line showing the correlation between text lengths, and the percentage of incorrect answers per task. There is a correlation with a correlation coefficient of 0.697.

We found an interesting correlation (correlation coefficient of 0.697) between bad comprehensibility, i.e., number of questions that were answered incorrectly, for text tasks and text length. The correlation can be seen in Figure 4.12. Since the shorter texts usually contained less numbers, it might be that they were easier to comprehend, because less numbers had to be compared. With shorter texts it was less likely that the wrong numbers were taken by accident. Aside from that the results did not show that a certain number type, e.g., percentages, decimal numbers, currencies, or heights was comprehended better or worse than others.

4.5.2 Processing Speed

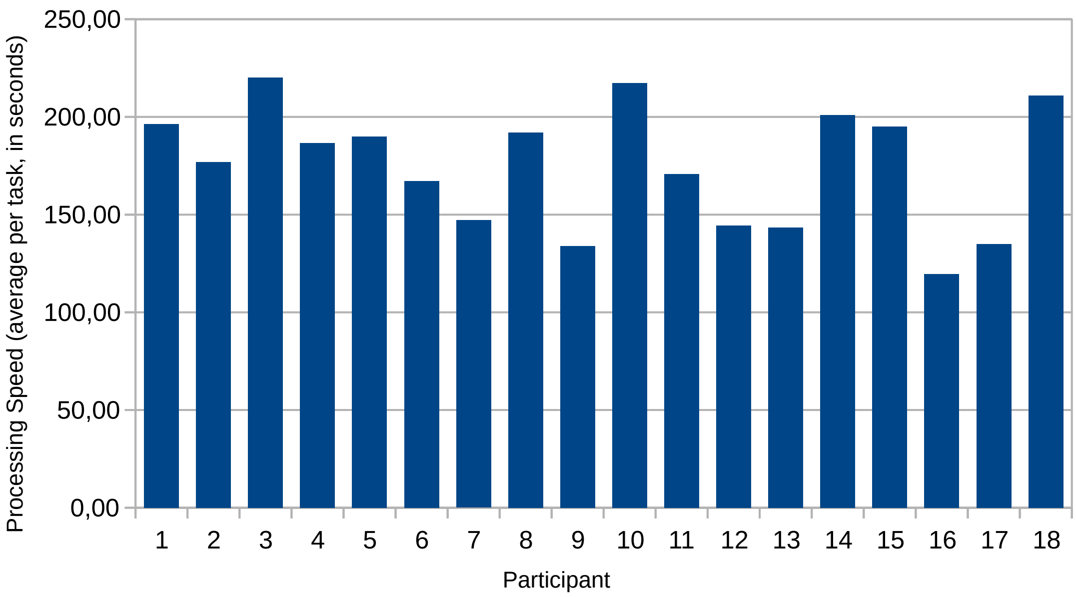


Figure 4.13: The average timings per task for looking at the visualizations or texts and answering the questions in seconds for the 18 participants. The fastest participant was participant number 16, with 718 seconds for six tasks, resulting in an average timing of 119.67 seconds per topic. The average timings per task were between 119.67 seconds (roughly 2 minutes) and 220.2 seconds (roughly 3 minutes and 40 seconds).

Timings, as stored by the study software, were used to evaluate processing speed. Since the participants solved a different number of tasks each, we divided their timings by the number of tasks they solved to get an average timing per task for each participant. Figure 4.13 shows these average timings for looking at the visualizations or texts and for answering the questions per task for the 18 participants. The fastest participant was participant number 16, with 718 seconds for six tasks, which means an average of roughly 2 minutes per task. The second fastest on average was participant number 9 with roughly 2 minutes and 14 seconds for each of the six tasks. The third fastest was participant number 17 with roughly 2 minutes and 15 seconds on average per task, while he/she saw all 15 topics at least once. In general, the participants needed on average between 2 minutes and 3 minutes and 40 seconds to solve a task.

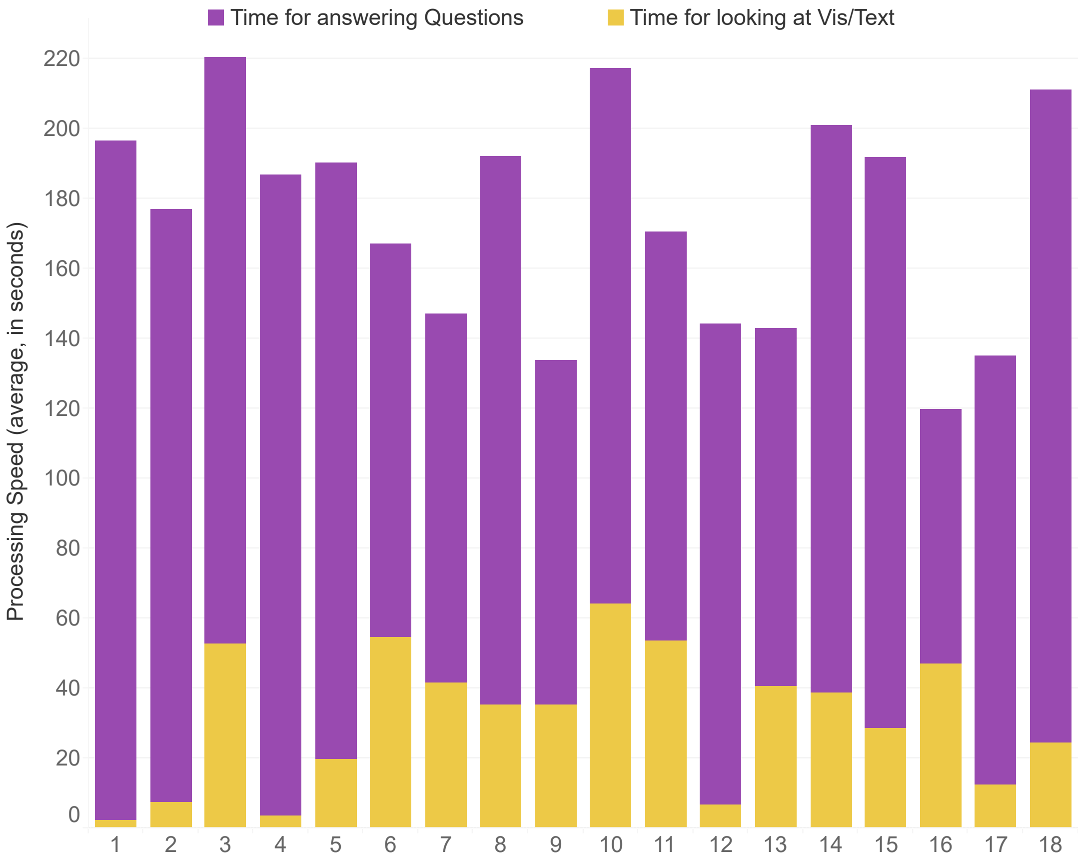


Figure 4.14: The number of seconds the 18 participants spent on average to look at the visualizations and texts (yellow), and to answer the questions (purple). Less time was used for looking at the materials than to answer the questions. Participants number 1 and 4 immediately concentrated on answering the questions without looking at the visualization or text before. Participant number 10 took the longest to look at the visualizations or read the texts before answering the questions.

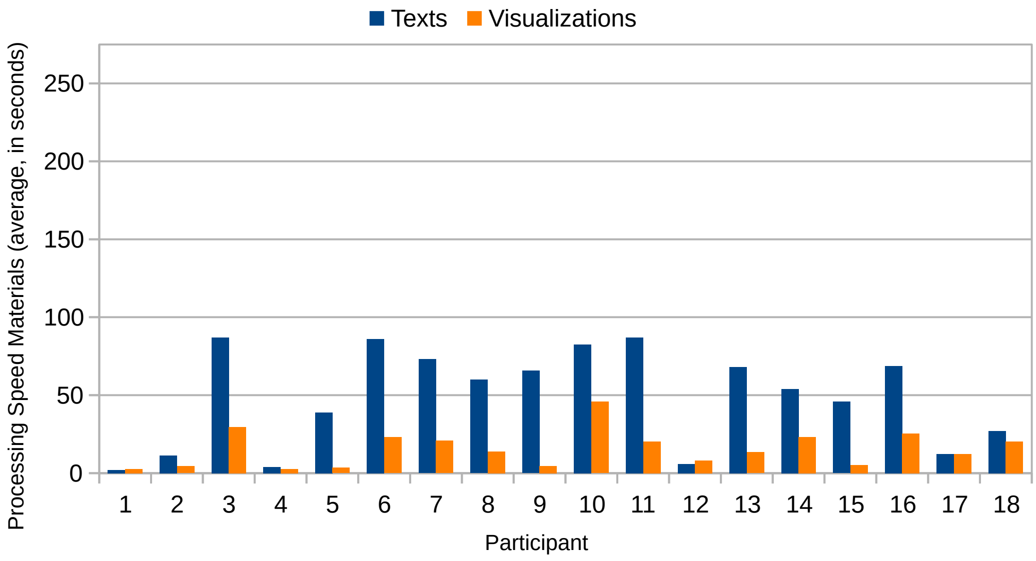
As stated above, our program took two time measurements. One was the time the participants took to look at the visualization or read the text, and the other one was the time they needed to answer the questions. Figure 4.14 shows these average measurements per participant. It can be seen that all of the participants spent much more time on answering the questions than on looking at the visualizations or reading the texts. Participant number 1 and 4 immediately concentrated on finding the answers to the questions without first reading the texts or studying the visualizations. Some participants also stated that they did not read the whole text, but skimmed the text, and only looked for certain words mentioned in the questions. When comparing comprehensibility with these strategies in mind (i.e., people first looking at the text/visualization or immediately pressing the button), no differences can be found. While talking to the participants after the study, some of them also stated that, in the beginning, their tactic was to look at the materials, show the questions when they were done, and try to answer them from their memory. But after some tasks they realized that this tactic did not help them and, therefore, they chose to show the questions immediately. Only participant number 12 looked longer at the visualizations than at the texts. Participant number 10 took the longest time of all participants to look at the materials. This participant also took a longer time to look at visualizations than the other participants, roughly 45.8 seconds on average per visualization task, and 82.6 seconds on average per text task. In total, participant number 3 took the longest to finish the study. He/she also stated after the study that he/she read all the texts, and enjoyed looking at the visualizations before answering the questions.

Figure 4.15 (a) shows how many seconds were spent on average per participant per task to look at the visualizations or texts without seeing the questions. Figure 4.15 (b) shows how many seconds were spent on average to answer the questions. In both diagrams blue denotes text tasks and orange visualization tasks. In the first stage, more time was needed for reading through a text ($M = 48.83$, $SD = 30.35$), than for looking at a visualization ($M = 15.53$, $SD = 11.29$). This is also a statistically significant difference ($p = 0.0003$).

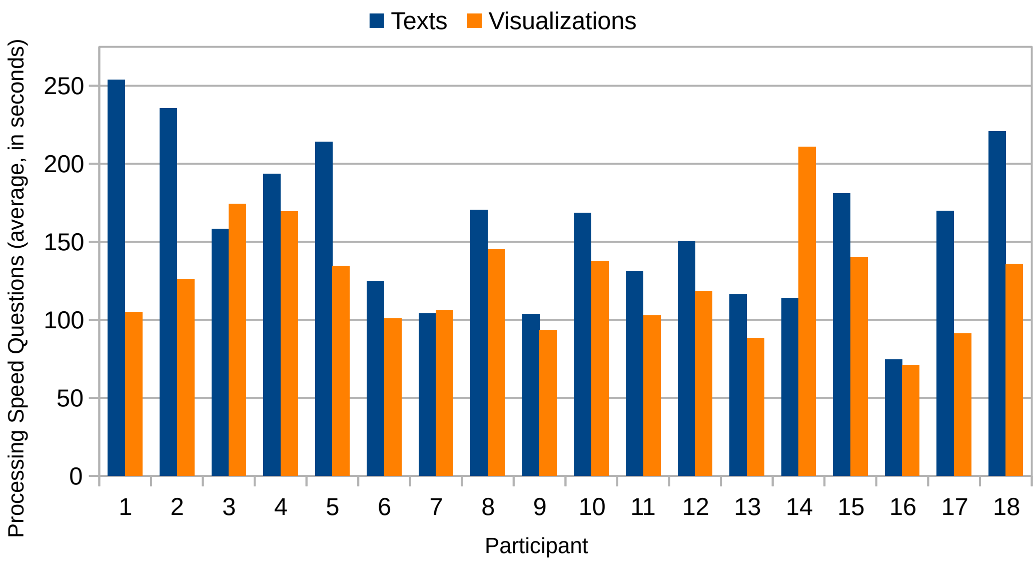
15 of 18 participants took more time to answer the questions when seeing a text ($M = 160.2$, $SD = 48.73$) than for looking at a visualization ($M = 125.05$, $SD = 34.19$). This difference is also statistically significant ($p = 0.02$). Especially participant number 1 took on average twice as much time for text tasks (253.78 seconds) than for visualization tasks (105 seconds). This participant also stated after the study, that he/she preferred visualizations over texts as he/she is a slow reader. Contrary, participant number 14 needed on average much longer to answer the questions for visualization tasks (210.75 seconds) than for text tasks (114 seconds). This participant stated that he/she preferred texts over the visualizations as he/she generally reads a lot and also fast. Participant number 16, who was the fastest in total, needed on average equally long for answering visualization task questions and text tasks questions, but also took longer to read the texts than looking at the visualizations.

On average, visualization tasks were solved faster than text tasks. To solve tasks with visualizations the participants needed on average 140.55 seconds ($M = 140.55$, $SD = 37{,}19$), while they needed on average 209.02 seconds for text tasks ($M = 209.02$, $SD = 36.42$). This is a statistically significant difference ($p = 0.000005$).

No statistically significant differences could be found for processing speed between male and female participants ($p = 0.46$), between participants having experience in visualization research and other professions ($p = 0.56$), between participants using German or English ($p = 0.72$), or between different age groups ($p = 0.5$).



(a)



(b)

Figure 4.15: (a) shows the number of seconds spent on average per participant per task on looking at the materials. (b) shows the time each participant spent on average on answering the questions. Blue are the text tasks, orange the visualization tasks. Participants that spent time on looking at the materials spent more time on reading the texts than looking at the visualizations. (b) shows that 15 of 18 participants took longer to answer the questions when they read texts as compared to when they saw visualizations.

We did find a statistically significant difference in processing speed of text tasks when the participants solved less than ten tasks or more than nine tasks (p = 0.04). Participants solving less tasks were, on average, significantly faster.

Topic	Texts	Visualizations	Diff. Vis to Text
1	195.00 s	143.80 s	-26.26 %
2	227.00 s	182.40 s	-19.65 %
3	174.67 s	115.90 s	-33.65 %
4	237.75 s	112.50 s	-52.68 %
5	152.33 s	70.80 s	-53.52 %
6	238.60 s	160.75 s	-32.63 %
7	243.80 s	101.33 s	-58.44 %
8	174.20 s	165.22 s	-5.15 %
9	247.20 s	254.83 s	3.09 %
10	205.29 s	71.00 s	-65.41 %
11	214.00 s	185.67 s	-13.24 %
12	137.83 s	129.88 s	-5.79 %
13	375.50 s	157.50 s	-58.06 %
14	201.11 s	190.50 s	-5.28 %
15	217.00 s	115.00 s	-47.00 %

Table 4.2: Average processing speeds per topic for visualizations, texts, and difference between visualizations and texts in percent. A negative difference (cyan) indicates that the visualization task was solved faster. A positive difference (magenta) indicates that the text task was solved faster. The topic with the highest average processing speeds for both, visualization and text, was topic number 9. It also had the smallest difference between average visualization and average text processing speeds. The visualization task that was solved fastest on average was for topic number 5. The text task that was solved fastest on average was for topic number 12. Topic number 10 had the biggest percentage difference between average processing speeds of visualization and text.

Table 4.2 and Figure 4.16 summarize the average processing times per topic for visualization tasks, for text tasks, and the time difference between visualizations and texts. A negative difference (cyan) in Table 4.2 means that the task of this topic was solved faster with the visualization, a positive one (magenta) means that it was solved faster with the text. Only topic number 9 was solved faster with text, but it also had the smallest difference in processing speeds between visualization and text task (3.09 %). Participants also spent the second most time for this topic, namely 251.36 seconds on average. Participants were only slower when answering questions for topic number 13, but mainly because it took them a lot of time to solve the text task. Topic number 13 also had the third highest difference between processing speeds of visualizations and

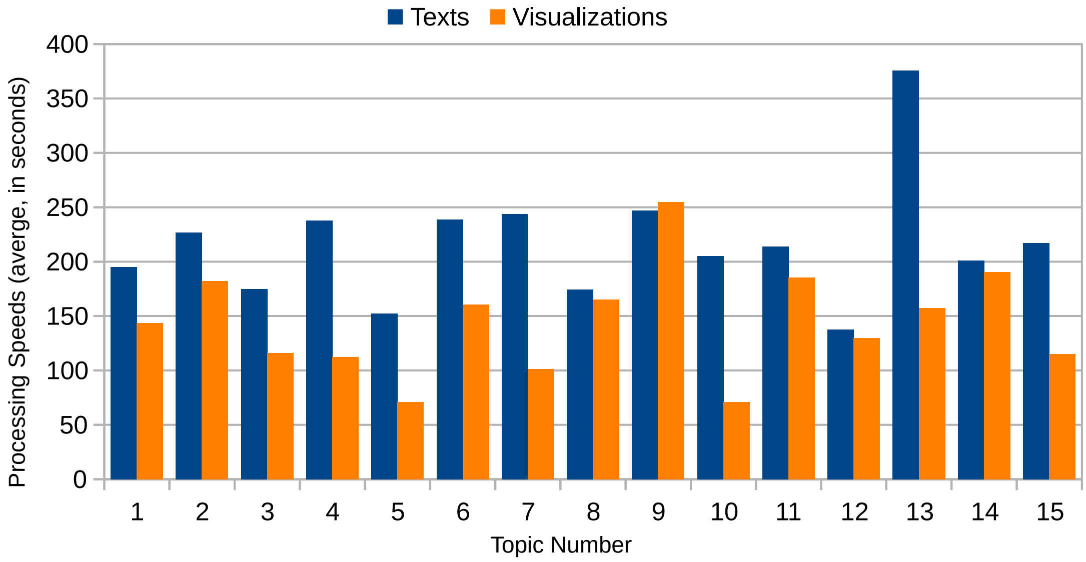


Figure 4.16: The average processing speeds per topic for text tasks (blue bars) and visualization tasks (orange bars) in seconds. It can be seen that in all but one case, i.e., topic number 9, visualizations were processed faster than texts. The biggest percentage difference between processing speeds of text and visualization tasks was measured for topic number 10. The percentage differences and average processing speeds can be seen in Table 4.2.

texts, i.e., -58.06 %. The highest average temporal differences between visualizations and texts were measured for topic number 10 (-65.41 %) and topic number 7 (-58,44 %). Both were solved faster with visualizations. Two other topics that were solved more than 50 % faster with visualizations than with texts were topic number 4 (-52.68 %) and topic number 5 (-53.52 %). The on average fastest processed texts were for topics number 12, 5, and 8 with 137.8 seconds to 174.2 seconds (roughly between 2 and 3 minutes). The texts for topics number 5 and 12 were also the two shortest texts. For topic number 9, the participants took by far the longest time on average with 254.83 seconds (roughly 4 minutes) for the visualization task, followed by topics number 14 and 11. By far the longest time for texts was spent on the text for topic number 13 with 375.5 seconds (roughly 6 minutes), followed by topic number 9 and 7. The texts for topics number 13 and 9 were also the two longest ones.

We also investigated the topics that were solved the slowest with visualizations or texts, which were topic number 9 with visualizations and number 13 with texts. For topic number 9 it has to be mentioned that, while it was the topic that was solved the slowest with a visualization, it was not solved much faster with the text, either. This can also be seen in Figure 4.16 and Table 4.2. Visualization number 9 was a chord diagram where participants had to compare and estimate many numbers. The texts for topics number 9 and 13 were also the two longest texts in the study, so it is no surprise that they took the longest to be processed. Further, both texts contained many numbers that had to be compared, and in case of text number 13, some calculations had to be performed.

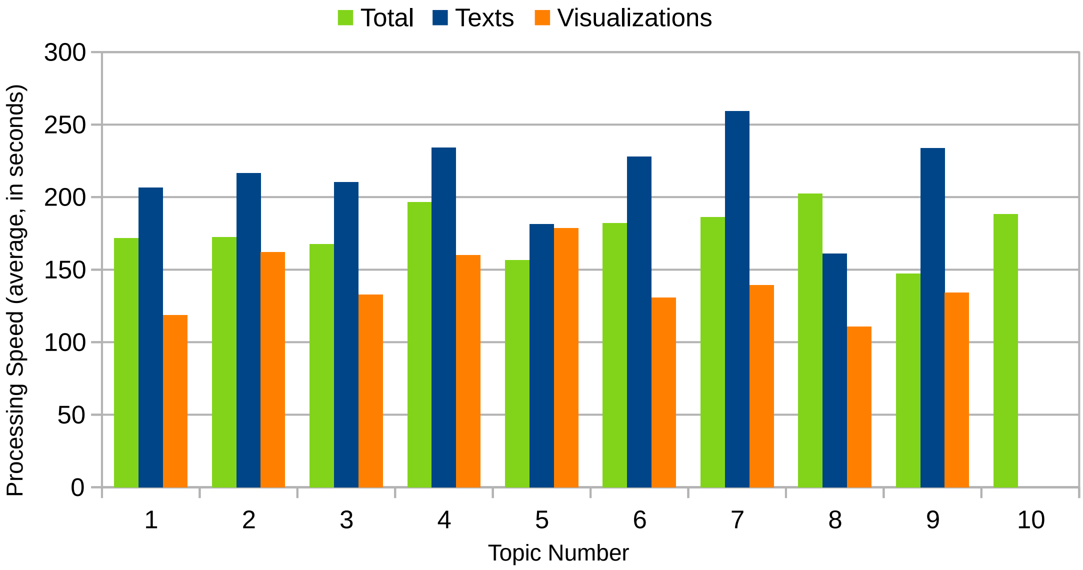


Figure 4.17: Average processing speeds per participant for their first to tenth task in total (green), for text tasks (blue), and visualization tasks (orange). Since the participants saw a maximum of nine visualizations or nine texts, there are no blue or orange bars for topic number ten. The average number of tasks that was solved by the participants was 10.7, so there are ten green bars. The order of the tasks did not influence the processing speed.

The order of the tasks did not have an effect on the processing speed. This can be seen in Figure 4.17. The bars showing processing speeds for visualizations, texts, and in total, had some ups and downs, but independently of the order of the tasks. It is more likely that the increases and decreases were caused by the difficulty of the visualizations or texts, and not, for example, because the participants could not concentrate that well anymore after solving a certain number of tasks.

We found a very interesting correlation between text lengths and processing speed. The correlation coefficient is 0.747, which suggests a strong correlation. This is not surprising, as it took participants longer to read longer texts. However, it also adds to the correlation between text lengths and wrong answers, as discussed in Section 4.5.1. A scatter plot showing the correlation between text processing speeds and text lengths can be seen in Figure 4.18.

We also investigated whether there is a correlation between processing speed and comprehensibility. The correlation coefficient is 0.03, which suggests that there is no such correlation. It cannot be stated that participants who took more time to solve the tasks, solved them more correctly, or that faster participants made more mistakes when answering the questions.

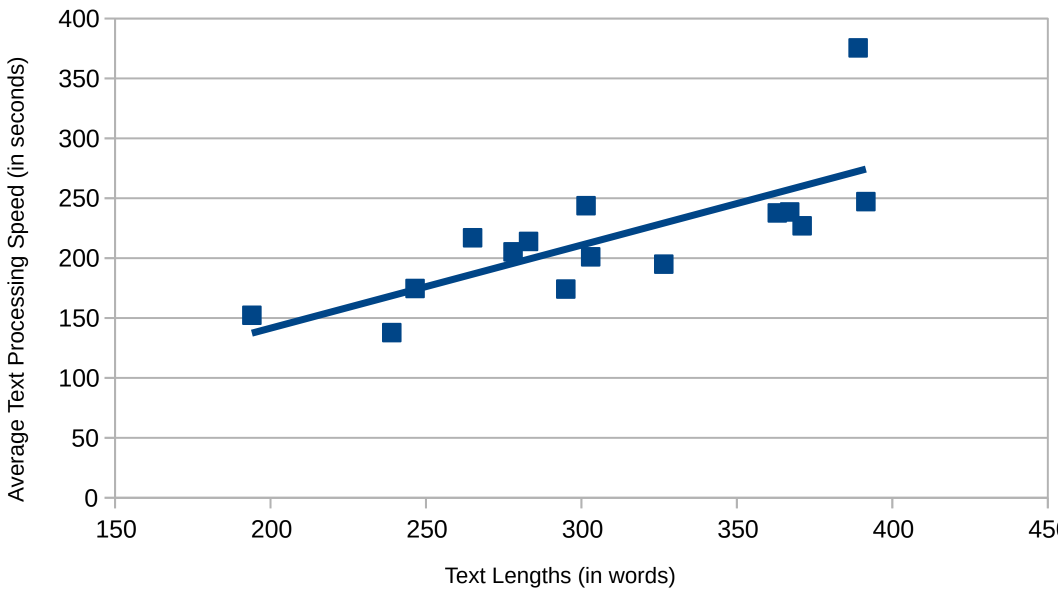


Figure 4.18: A scatter plot with regression line displaying the correlation between text processing speeds and text lengths for each task. There is a correlation with a correlation coefficient of 0.747.

4.5.3 Post-Study Questions and Summary

The majority of the participants (14 out of 18) thought they performed better when solving visualization tasks, especially concerning the processing speeds. Another three participants stated that they felt like both, visualizations and texts, worked equally well for them. One participant stated that he/she felt like he/she gave more accurate answers for questions where he/she had to read the text. This is most probably due to the fact that for some questions participants had to estimate values, while they were written clearly in the text. In general, people found visualizations without a grid, or lines that helped to estimate the correct value, more difficult. One participant also stated that he/she thought that it was easy and fast to extract information from simple visualizations, but for visualizations that contained much information, he/she felt like these visualizations were not really easier to comprehend than the texts. This participant also stated that it would be good to have visualization and text at the same time. Another participant told us that usually texts would not be written like in our study, i.e., containing so many different values. In most cases, there would be a corresponding visualization in addition to such a text.

Concerning questions that were perceived as more difficult than others, the participants mentioned longer questions, questions where a value had to be calculated (e.g., differences between two values), and questions to retrieve a certain value, while watching out for multiple constraints. One participant mentioned that the questions that involved some sort of calculations were much easier to solve with visualizations. It was also mentioned that questions where we did not, for example, ask about the highest value but third highest value were perceived as more difficult. Some participants told us that it was more difficult for them, if the wording of the question was different than in the text, because in this case they could not skim the text for the words that occurred in the question.

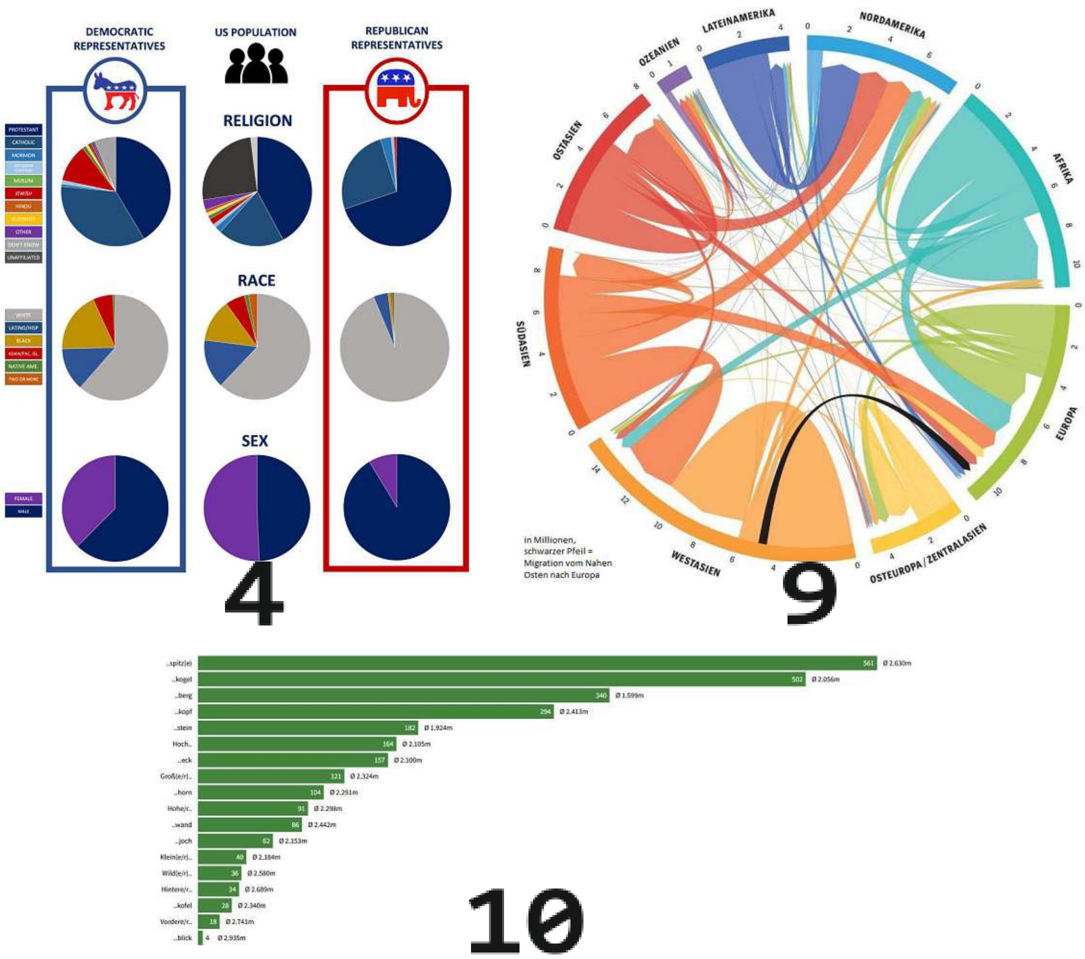


Figure 4.19: The visualizations of the three topics with the highest percentages of incorrectly answered questions in descending order. For the visualization of topic number 4, participants had problems with estimating percentages of the pie segments and with comparing different pie charts. Participants had problems with estimating numbers and comparing the width of the arrows and segments when working with the visualization of topic number 9. For the visualization of topic number 10, there seemed to be a slight problem with a certain question where participants confused the information displayed on the left side of the bars with the information displayed on the right side.

Participants mentioned bar charts and pie charts to be easier understandable than more complex representations. Interestingly, despite this fact, the visualizations for topics number 4 and 10, which contained pie charts and bar charts, can be found among the visualizations with the highest number of incorrect answers (shown in Figure 4.19). For the visualization of topic number 4, the problems for most participants were the missing percentage values, so participants had to guess them from the shape of the segments. Another problem seemed to be that there were multiple pie charts, and in some cases the participants looked at the wrong one. The segments of the pie charts were not ordered according to the percentages they displayed, instead, they always represented the same order.

The visualization of topic number 9 was about global migration and described how many people moved from one continent to another, or moved within the same continent. The visualization of topic number 9 is a chord diagram, which is less common, and people do not come across such visualizations that often. The participants also had problems with estimating how many people moved between or within continents by looking at the width of the arrows. It also seemed like the participants had problems with comparisons of numbers between the continents, maybe due to the fact that it was a circular arrangement.

The biggest problem with the visualization of topic number 10 was that it contained two different types of information. On the one hand, it showed how many Austrian mountains share a certain name part. The name part was written on the left and the length of the bars indicated how many there were. This was also indicated by numbers in the bars. On the other hand, it contained the average heights of the mountains that share a name part. They were written directly to the right of the bars. Therefore, when we asked for certain heights, some participants wrongly assumed that the bar lengths indicated the average heights, and thought that the longest bar belonged to the highest mountains when in reality it belonged to the most common name part. Four out of five participants answered this question wrong.

Even though the visualizations of topics number 4 and 10 belonged to the visualization tasks for which people gave the highest number of incorrect answers, they also belonged to the five visualization tasks that were processed more than 50 % faster than their textual counterparts.

The visualization of topic number 9, the chord diagram, is also among the visualization tasks that took participants the longest to interpret (shown in Figure 4.20). The other two visualization tasks that were solved slowest were for the visualizations of topics number 14 and 11. The visualization of topic number 14 was a stacked bar chart describing the distribution of people living and working in a Viennese district, or in a neighboring district. Comprehensibility for this visualization was good with many participants giving correct answers, but it seems that estimating the percentages of some bars on the scale on the left side took them some time. Maybe the fact that the districts were not ordered by district numbers, but by how many people worked at home and in the same district combined, made it more difficult to find the values for certain districts.

The visualization of topic number 11 was a population pyramid, i.e., a form of comparative horizontal bar chart that shows how many men and women of different age groups live

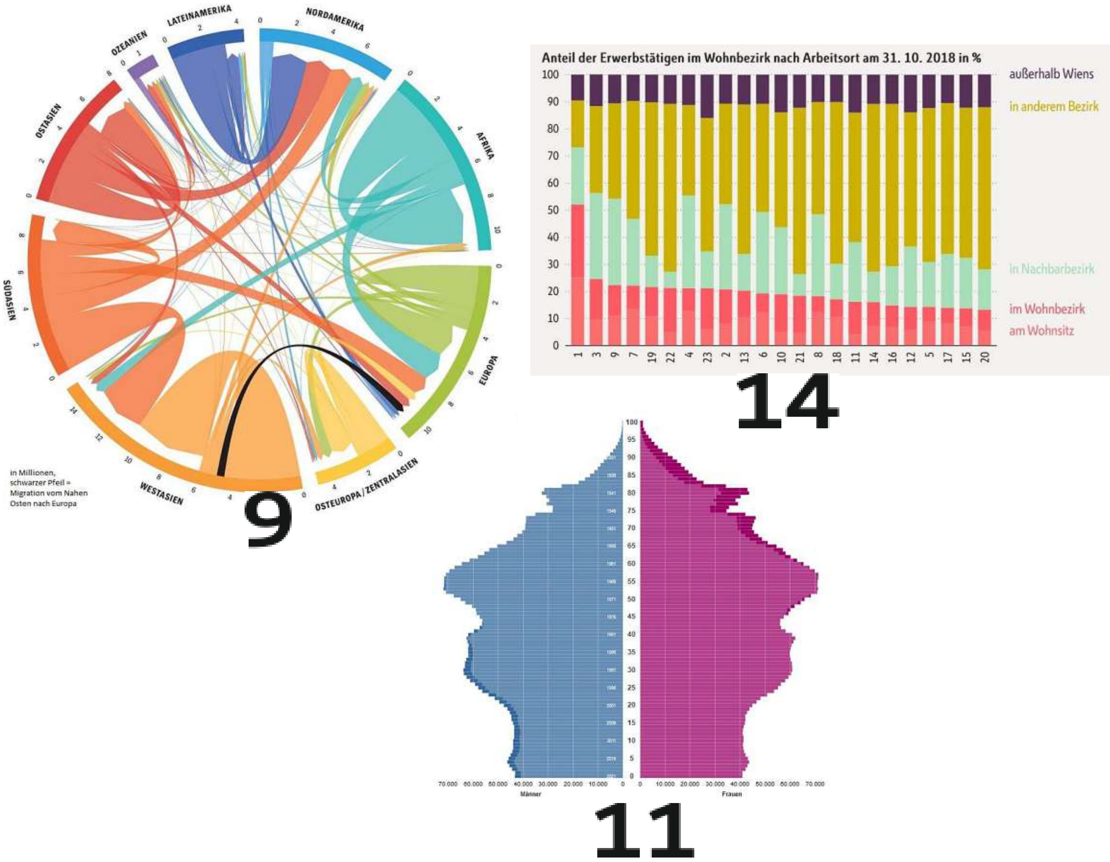


Figure 4.20: The three visualizations of topics where the visualization tasks were solved the slowest, shown in descending order with the slowest on the left. The visualization of topic number 9 is a chord diagram which is not so common, so it might have taken the participants time to get accustomed to it. The visualizations of topics number 14 and 11 are different kinds of bar charts containing many bars. All three visualizations display much information that also had to be compared in order to answer the questions.

in a certain country (in this case Austria). In general, the participants comprehended this visualization quite well and did not make mistakes when answering the questions. However, some of the participants made bad estimations of some values, since the population scale only had marks at steps of 10,000, and the scale depicting the years only had marks every 5 years. The visualization contained much information since it portrayed the last 100 years, meaning that there were 100 bars for men and also for women that had to be compared for some of our questions. For one of the questions the participants had to add the male and female population of a certain year, which might have been more difficult due to the auxiliary calculation.

In Figure 4.21 the topics where the visualization tasks were processed at least 50 % faster are shown. The visualization of topic number 4 has been discussed before, it is

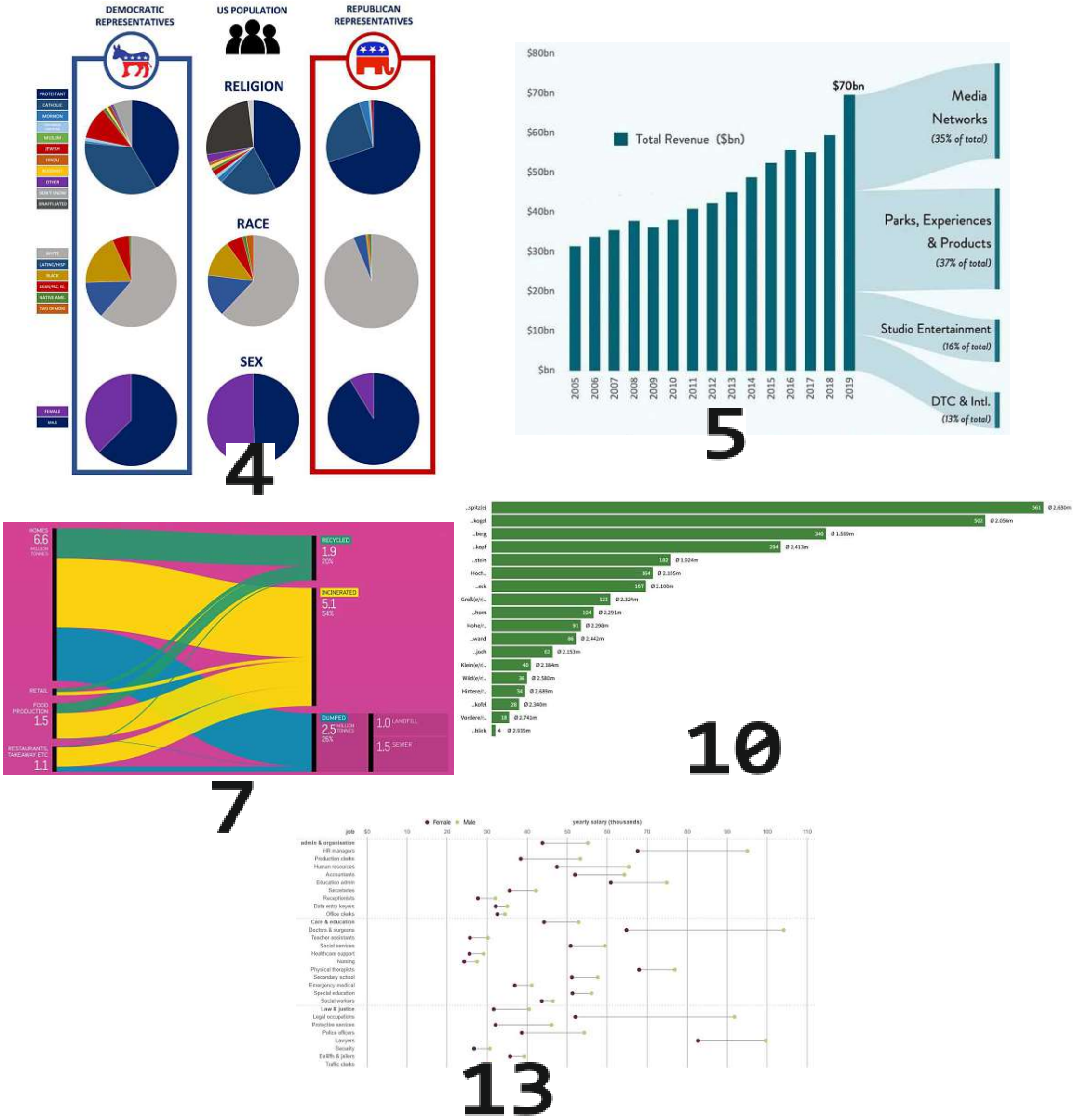


Figure 4.21: The five topics where visualization tasks were processed more than 50 % faster than text tasks. The visualizations of topics number 4 and 10 also belonged to the visualizations with the highest number of incorrectly answered questions. Contrarily, participants made no mistakes at all when answering the questions for the visualization task for topic number 5. For the visualization of topic number 13, participants struggled with the text, as it was the task where participants worked the slowest and had the highest number of incorrect answers. For the visualization of topic number 7 people gave roughly as many incorrect answers in the visualization task as in the text task, but the visualization was processed much faster.

among the visualizations with the most incorrect answers. It could be that participants underestimated the complexity of the visualization. The visualization of topic number 5 is a Sankey diagram connected to a bar chart. The Sankey elements not being in descending order did not distract the participants, neither in terms of processing speed nor comprehensibility. When solving the visualization task, participants made no mistakes at all, which makes the visualization of topic number 5 the best comprehended visualization of all. The visualization of topic number 7 is also a Sankey diagram displaying food waste in the UK, showing sources for food waste and where the food gets discarded. Participants answered roughly equally many questions correctly for the visualization tasks and the text tasks. The visualization of topic number 10 has already been discussed as being among the visualizations with the most incorrect answers (similar to the visualization of topic number 4). The visualization of topic number 13 shows the pay gap between men and women in the US. In this case the text was the second longest in the study. The visualization of topic number 13 also had the biggest differences among comprehensibility for visualization and text tasks, so it seems participants had problems comprehending the text. This text was also mentioned by some of the participants as a difficult one, because it contained so many different job titles and numbers.

Figure 4.22 shows the topics where the visualization tasks had the highest number of correct answers. The visualization of topic number 5 has already been mentioned before with being the one with the least mistakes made in answering the questions. Surprisingly, the visualization of topic number 2 was also one of the best in terms of comprehensibility. It was a coropleth map about unemployment rates of different German states, but also showed the gross domestic product (GDP) for two different years and the population for two different years for all of the states.

The fact that the visualization of topic number 1 about cigarette consumption was understood that well was also a surprise, since it contained two types of information, which could be misleading or distracting for the participants. The horizontal axis denoted how many percent of men, and women from different countries smoke cigarettes daily. On the right side it was written how many cigarettes men, and women of different countries smoke on average per day. During our post-study talk some participants mentioned that, at first, they did not realize that these were two different informations.

When talking about visualizations, or visualization types the participants liked best, participants mostly chose some of the more common types, like bar charts or pie charts, or very colorful visualizations like the visualization of topic number 8 (Immigration to the U.S.), and the visualization of topic number 9 (Global migration).

Only 4 of the 18 participants stated that they do not like to read normally, and just read texts like newspaper articles. One participant mentioned that books can be read faster than the texts in our study, because not every word is important in books, but the texts from the study had to be read very thoroughly in order to be able to answer the questions correctly.

Figure 4.22: The three topics where the visualization tasks had the highest number of correct answers in descending order from left to right. Even though all of the visualizations displayed at least two different types of information, participants comprehended them very well. The visualization of topic number 5 had one type of information displayed with bar charts and one type in a Sankey diagram. The visualization of topic number 2 contained three types of information, two of them displayed in red and blue boxes, and one displayed with colored states on the map. The visualization of topic number 1 displayed one type of information in the middle and one on the right side.

When we asked the participants if they found the visualizations or texts more interesting, 13 of the 18 participants said that they found the visualizations to be more interesting.

One of the participants justified their preference by saying that a visualization not only shows the information, but also has an appealing graphical part on top of this. Another participant stated that he/she found the visualizations more interesting, but also harder to understand than the texts. The other five participants said that they found both texts, and visualizations equally interesting. Concerning the topics, the participants liked best, their answers were very diverse, and naturally often corresponding to their interests. For example, one participant was a smoker, so he/she stated that he/she found the text about cigarette consumption (topic number 1) most interesting, while another participant liked hiking and, therefore, found the topic about Austria's mountains (topic number 10) most interesting. Easier to grasp topics like the mountain names or Christmas gifts were preferred over more serious topics like inflation or elections. Participants preferred short texts over longer ones, independent of the topic.

Six of 18 participants told us that they checked their answers before saving them, another six stated that they did not check them, and the remaining six said that they only checked them in some cases, for example, if they were not sure about their answer. Most participants also solved the questions not in the order they were presented, but in their preferred order. The majority of participants also did not take breaks between the tasks.

4.5.4 Hypotheses Creation

For Research Question 1, we wanted to find out if humans can process visualizations or texts faster, and also how much faster. A widespread claim exists stating that visuals are processed 60,000 times faster than text, which made us assume that visualizations will be processed faster than text. This was confirmed in the study, where visualization tasks were solved on average 33.4 % faster than text tasks. The assumption that visualizations will be processed between 6 and 600 times faster has, however, been tuned down to about 1.3 times faster. From these results, we, therefore, came up with the following hypothesis:
Hypothesis 1: Visualizations are processed at least 1.3 faster than text.

For Research Question 2, we wanted to find out whether humans better comprehend information if presented as visualizations or as text. We also wanted to find out if so, how much participants would perform better. Previous results indicate that adding meaningful visualizations to texts increases comprehensibility. Little information was found so far on a direct comparison between visualization and text. Since participants were used to both representations, we assumed that visualizations and texts will perform equally well in terms of comprehensibility. The results of our exploratory study show that comprehensibility for texts was only statistically insignificantly better than for visualizations, which confirms our assumptions. We, therefore, came up with the following hypothesis:
Hypothesis 2: Visualizations and texts can be comprehended equally well.

For Research Question 3, we wanted to find out whether the content (i.e., the topics) made a difference with respect to comprehensibility or processing speed. Studies so

far did not really pay attention to this question. We assumed that the content will affect how well participants can solve the tasks, due to prior knowledge of participants with different educational background. After checking the visualizations with most correct and incorrect answers, we cannot see a correlation here. Participants made mistakes or performed very well independent of the topic or content that was presented in the visualizations. We, therefore, came up with the following hypothesis:
Hypothesis 3: The content does not affect processing speed and/or comprehensibility of visualizations.

4.5.5 Limitations

Our study has some limitations, as not all possible research questions could be addressed in one user study.

After evaluating the results, we assume that a better selection of visualization and text tasks for each topic would have been beneficial. In our case we had some topics where there was an imbalance between visualization and text tasks.

It would be interesting to include more visualization types into another study. Also clustering the visualizations before conducting a study (e.g., simple ones, more complex ones) could help in interpreting the results.

Our timings measured the total time the participants needed to answer all questions of one task. With timings that measure the time for each question separately, it would be easier to find out which questions were harder to solve for the participants. We also did not really benefit from our split time measurements for looking at the materials and answering the questions, since some participants decided to look at the questions immediately.

We realized that the focus of the participants in this study was to answer the three questions per task. This does not mean that the participants read or comprehended the whole texts or visualizations in a certain amount of time. Our study only states the results for the questions we asked the participants about. Especially the processing speeds for texts cannot be fully interpreted as the times the participants needed to read the texts, because many of them did not read the texts, but skimmed them for certain words in the questions. For further studies it would, therefore, be recommended to let participants explain the content of the visualization/text with their own words.

CHAPTER 5

Conclusion and Future Work

Data visualization plays an important role in different application areas and our everyday life. In this thesis, we were interested in studying how well humans perform when reading visualizations, in comparison when they have to read text. We were especially interested in processing speed and comprehensibility.

A widespread claim is that visuals are processed 60,000 times faster than texts, but until today it is unclear what the source for this statement is. It was also contested in an article in Section 2.5. In Sections 2.6 and 2.5 we analyzed existing studies about processing speed and comprehensibility of visualizations and texts. However, none of the studies about processing speed directly compared visualizations to texts. Most of the studies about comprehensibility compared text to text with additional visualizations, so the effects of having only a visualization were unclear. The ones that directly compared visualizations to texts were inconclusive or conflictive. Although the existing studies contained some interesting results, e.g., that text benefits from additional meaningful visualizations in terms of comprehensibility, it was still unclear if visualizations or texts are better comprehensible and which ones are processed faster. Therefore, we conducted an exploratory study to learn more about these questions.

The idea of our study was to show the participants either a visualization or a text containing the same information. We chose 15 visualizations from various sources on the internet. The question was, how to get a text with the same information as a visualization. In order to be able to write such texts, we had a look into the topic of visual attention, which is about finding the areas in an image that get the most attention by the viewers. We discussed studies about this topic in Section 2.3. The saliency models presented in this section turned out to work better with natural images or favored textual parts. Other studies showed that people focus on extrema, title of a visualization, or textual parts in general. Apart from these insights, we wanted to find out which parts of the visualizations we chose for our study would get the most attention and should be mentioned in the text.

To achieve this, we conducted an online pre-study in which five participants were randomly shown four out of the 15 visualizations each. The Participants had to write a text describing each of the four visualizations and also rate their confidence in their description. The results of the pre-study confirmed those from the studies from the state of the art and showed that our participants also mostly mentioned extrema, the general topic of the visualization, and comparisons between certain values. Using these results, we created the texts for our study while also paying attention that the text length resembled the complexity of the visualization.

For our exploratory study we recruited 18 participants that were each shown at least six different topics out of 15. The topics were presented on a computer, each one either as text or visualization. The participants were able to decide how long they wanted to look at the visualization or read the text before seeing and answering the questions. During the study, the participants were watched in order to be able to determine if there were any disturbances. After the study, we talked to the participants and asked them some questions about the study, e.g., if they had any problems or which visualizations they found easier than others.

5.1 Contributions

The results showed that visualizations were processed roughly 1.3 times faster than texts, which is a big difference to the claim that they are processed 60,000 times faster. Concerning comprehensibility, we came to the same conclusion as previous studies from the state of the art, that visualizations are comprehended equally well as texts. We also found strong correlations between text lengths and processing speeds of texts, as well as text lengths and comprehensibility of texts. This means that our participants needed more time to solve tasks with longer texts than with shorter ones and also comprehended the shorter texts better. While it is not surprising that shorter texts were processed faster than longer ones, we think that the shorter texts were better comprehensible because they contained less numbers which had to be compared or used for calculations. We did not find a correlation between processing speed and comprehensibility, meaning that participants did not make more mistakes when they solved a task faster, but also did not make less mistakes if they took more time to solve a task. Interestingly, we found out that men answered significantly more questions with texts correctly as compared to women. While the contents of the visualizations or texts did not seem to play a role, we found some cases in which visualizations or texts clearly outperformed the other one in terms of processing speed or comprehensibility. Participants also liked shorter texts better than longer ones and preferred especially colorful visualizations or more common visualization types like bar charts or pie charts. In general, participants liked topics better if they corresponded to their interests.

Our exploratory study led to new hypotheses to be tested in future studies. We now assume that visualization can be processed about 1.3 faster than text. We further assume that information is equally well extracted from text as from visualizations. It also seems that the content of the tasks does not have an influence on task solving, as long as the content is easy enough to be quickly understood by the participants.

5.2 Limitations

It would have been helpful to measure the time for each question separately, in contrast to our approach, where we only measured the total time the participants needed to answer all three questions per topic. With separate timings it would be possible to tell if people needed more time to answer or understand certain questions and identify possible problems or difficulties. This could either be done by displaying the questions separately after a click on a button or by using an eye-tracking device. By displaying only one question at a time, the results would clearly be only for this question. However, it would take away the freedom of the participants to choose the order in which they answer the questions, which could also be interesting to find out which questions were perceived to be more or less difficult than others. Using an eye-tracking device would clearly show the areas on which the participants focus in the visualizations or texts in order to answer the questions. This could also lead to some interesting insights about important parts of visualizations or texts. It would also be possible to see which tactics the participants used to find the answers, e.g., if they read the whole text, certain parts of it, or only skimmed it for numbers or certain words. For questions that were answered incorrectly by a participant, eye-tracking could show what the problem was, e.g., if participants looked at the wrong number or area, if they looked at the right area and only made a bad estimation, or if they looked at the correct numbers, but miscalculated the result.

Another interesting approach would be to further examine the correlations between text length and comprehensibility and text length and processing speed that we found in our study. While we could not find an influence of certain types of numbers, it is possible, that there are other text attributes that had an influence which we did not examine in our study.

Our study only took a small sample of visualization types and there are still many more types to analyze and get results for processing speed and comprehensibility. A study with further visualization types could also include interactive visualizations or simulations. It would also be possible to display the same information with different visualization types and see which type was better comprehensible and processed faster in comparison to the other ones. Like this, it could be deduced which visualization types are suited best for which information types or if there are certain visualization types that work very well in general. It would also be interesting if our results, i.e., bar charts or pie charts were perceived as better comprehensible, but often underestimated concerning their complexity, could be reproduced.

5.3 Future Work

More research could be done concerning the question of when does a text contain the same information as a visualization. It would help to be able to write good alternative texts for visualizations that do not focus on certain parts of visualization types like axis labels or legends, but instead focus on the content of the visualization and communicate the most important parts of it. Ideally, the most important parts could be obtained automatically and maybe also text blocks could be derived from them automatically.

For a future study, it would be also interesting to measure the visualization literacy of the test persons, see how well they perform for different visualization types in the study, and find out, if there is a correlation between their performance and their visualization literacy. It would also be possible to see which visualizations worked better for people with higher or lower visualization literacy.

APPENDIX A

Materials used in the Study

A.1 Topic number 1: Cigarette use 2010

A.1.1 Text

In 2010 there was a survey about cigarette use of people older than 15 years in 15 different countries all around the world. The results showed, that in all of these countries more men than women smoked at least one cigarette daily. Also, in all but one country, the men smoked more cigarettes a day compared to women. The exception was India, where women smoked on average 7 cigarettes a day compared to men with 6.1 cigarettes per day. The country with the highest percentage of men, who smoked daily, was Russia, where nearly 55% of men and roughly 17% of women smoked daily. On average, Russian men smoked 18.5 cigarettes a day and women 12.7. The second highest percentage of men who smoked daily were Ukrainians with 45%, while only 9% of Ukrainian women smoked. On average 18.1 daily cigarettes were smoked by male and 11.8 by female Ukrainians. With roughly 44% of daily male smokers, China got the third place. However, only roughly 2% of Chinese women smoked daily, but compared with the other top 3 daily smoker countries, Chinese women smoked the most with 12.8 cigarettes per day, while the men smoked the least compared to the others with 16.6 cigarettes a day. The least number of daily smokers were found in India (roughly 6% of men and 1% of women), Mexico (roughly 11% of men and 3% of women) and the United States (roughly 16% of men and 13% of women). India and Mexico also had the lowest number of cigarettes smoked per day on average, while men in the United States smoked 18.2 and women 15.2 cigarettes. For women, this was the second highest number of daily cigarettes after Poland, where they smoke 15.5 daily. Also, the United States had the third highest percentage of women who smoked daily after Russia and Uruguay which both had roughly 16%. On average the men from Egypt smoked most cigarettes daily (19.4) followed by Turkey (19.3) and Russia (18.5).

A.1.2 Visualization

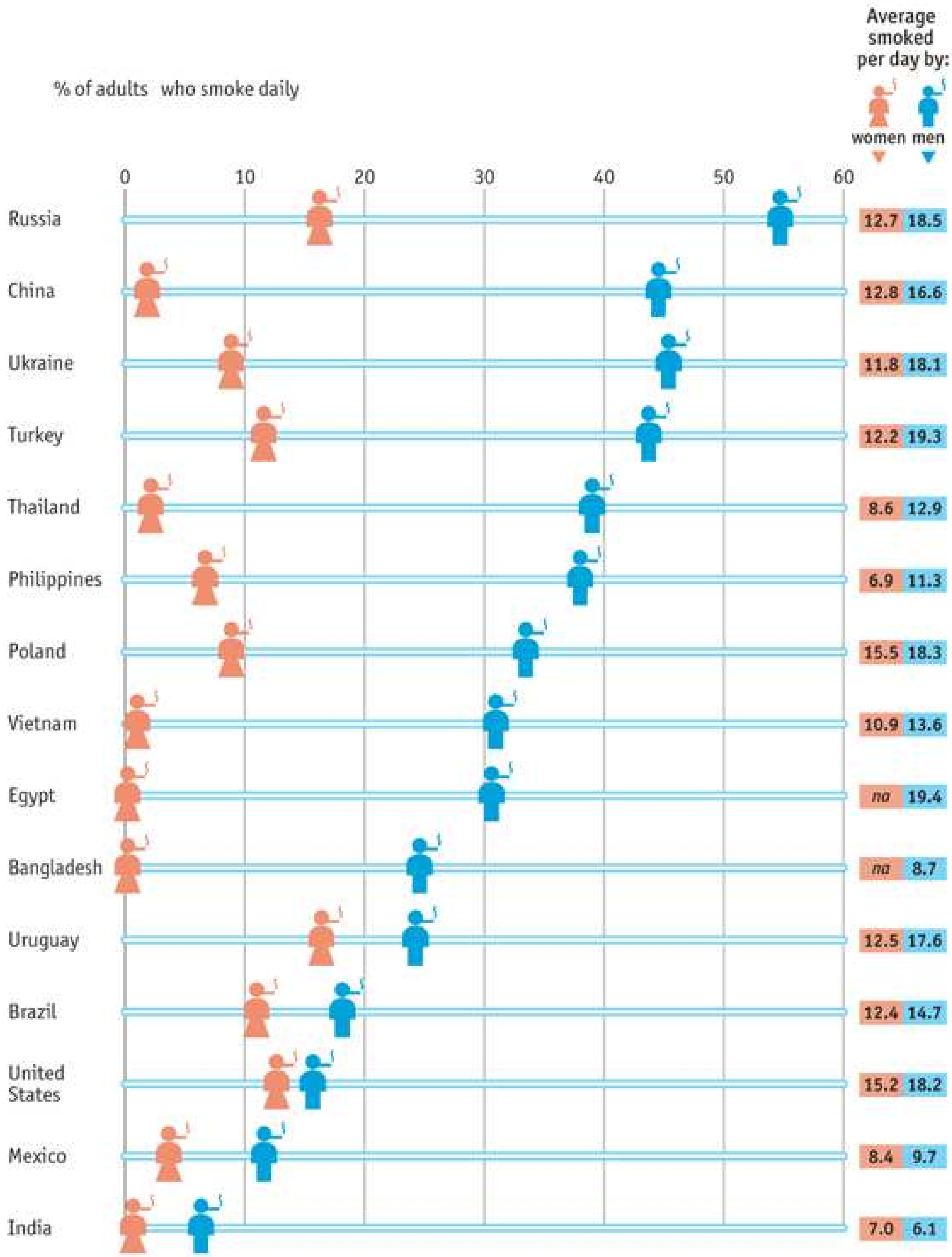


Figure A.1: Visualization for topic number 1 [V1].

A.1.3 Questions

1) In which country did women smoke the most cigarettes per day on average?
2) Approximately what percentage of men smoked daily in the U.S.?
3) Was there a country where women smoked on average more cigarettes per day than men? If so, which one?

A.2 Topic number 2: Unemployment rate Germany 2010

A.2.1 Text

In 2010 the unemployment rate in Germany differed vastly from state to state. Especially the states that formerly belonged to East Germany had high rates with over 12% in Berlin, 10-12% in Mecklenburg-West Pomerania, Saxony-Anhalt and Saxony and 7-9.9% in Brandenburg and Thuringia. States from formerly West Germany that had higher unemployment rates were Bremen with 10-12% and Hamburg, Lower Saxony, North Rhine-Westphalia and Saarland with 7-9.9%. The states with the lowest unemployment rate were Bavaria and Baden-Württemberg with under 5%. Schleswig-Holstein, Hesse and Rhineland-Palatinate had a rate between 5% and 6.9%. Also, when looking at the GDP per person in 1991, the three states with the lowest GDP all were from former East Germany, namely: Thuringia (9.7k Euro), Mecklenburg-West Pomerania (10.8k Euro) and Saxony-Anhalt (10.9k Euro). The states with the highest GDP per person in 1991 were Hamburg (43.8k Euro), Bremen (33.8k Euro) and Hesse (31.1k Euro). In general, back then all formerly West German states had a GDP per person of at least 23.5k Euro. Looking at the GDP per person in 2009 not much changed in the formerly West German states, where the top 3 were still Hamburg (48.2k), Bremen (40.5k) and Hesse (35.7k) and all the other states had at least a GDP per person of 25.5k. Not much changed in the formerly East German states either, but compared to 1991 the GDP there rose for about 10k Euro, except Berlin, where it was already higher than in the other formerly East German states before. So the states with the lowest GDP per person now were Mecklenburg-West Pomerania (21.3k), Brandenburg (21.4k) and Saxony-Anhalt and Thuringia (both 21.7k). Population-wise there were no big changes between 1991 and 2009 in most states. Only the two southern states Bavaria and Baden-Württemberg had bigger changes in population, where it rose for about 1 million in Bavaria and about 0.8 million in Baden-Württemberg. The state with the biggest population in 1991 as well as in 2009 was North Rhine-Westphalia with 17.4 million and 17.9 million respectively. Comparing the populations from 1991 and 2009 it can also be seen that, while the populations rose in most formerly West German states, they shrank in all the formerly East German states except Berlin, where they stayed roughly the same.

A.2.2 Visualization

Figure A.2: Visualization for topic number 2 [V2].

A.2.3 Questions

1) By how much did the GDP per person in Bremen increase between 1991 and 2009?
2) In which state did the population increase the most between 1991 and 2009?
3) Which state had the highest unemployment rate in 2010?

A.3 Topic number 3: Money spent on Christmas gifts 2011

A.3.1 Text

Nowadays, much money is spent on Christmas gifts. In a survey from 2011 people from 20 different countries all over the world were asked how much they plan to spend for Christmas presents that year. People from Luxembourg wanted to spend over 750$. However, it has to be said, that the GDP per person was also very high in Luxembourg with roughly 85,000$, which was also the highest of the participating countries. The second and third most generous countries were the United States with a bit more than 700$, while they also had the second highest GDP per person of the participating countries with nearly 50,000$, and Ireland, where people planned to spend roughly 700$, while they had a GDP per person of roughly 40,000$. People from Switzerland, the participating country with the third highest GDP per person of roughly 43,000$ planned to spend over 650$ for presents, which put them directly behind Ireland concerning the dedicated money for gifts. The least money for gifts was spent by people from the Netherlands with roughly 180$, while they had the fourth highest GDP per person of roughly 41,000$. Two other countries with the least spending for gifts were Greece and Ukraine, where people wanted to spend a little bit under 250$. However, of the participating countries, Ukraine also had the smallest GDP per person of roughly only 8,000$, while it was nearly 30,000$ in Greece. People from South Africa, the country with the second lowest GDP per person (roughly 11,000$) wanted to spend roughly 300$ on gifts.

A.3.2 Visualization

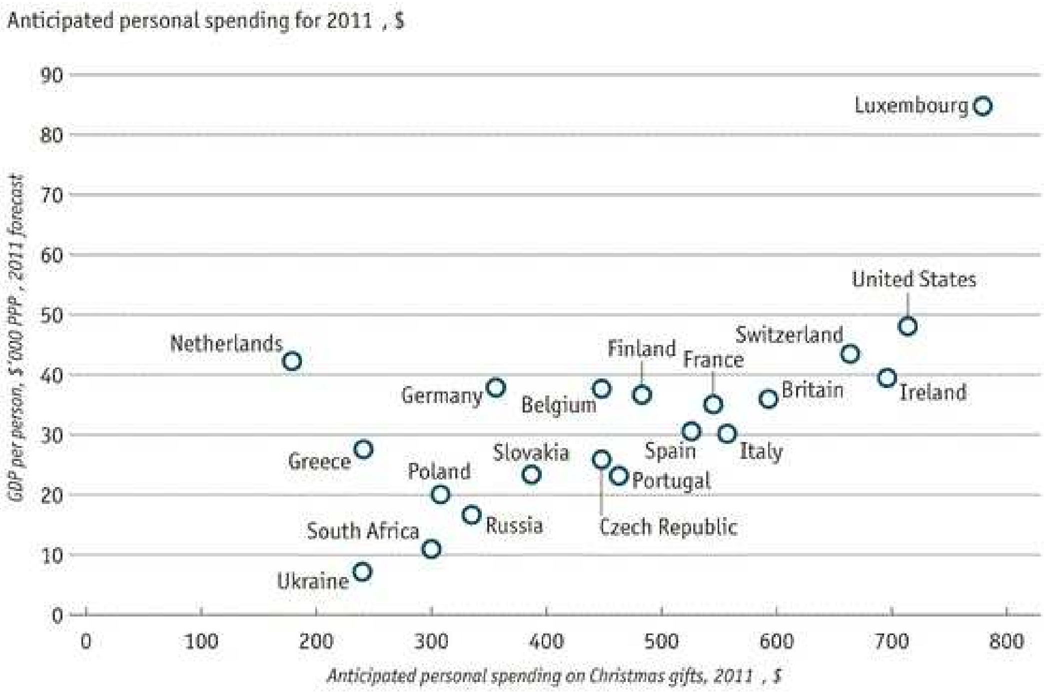


Figure A.3: Visualization for topic number 3 [V3].

A.3.3 Questions

1) In which country did people plan to spend the least money on Christmas gifts?
2) What was the GDP per person in the country where people wanted to spend the third most on Christmas presents?
3) How much money on gifts did people from the country with the second lowest GDP per person want to spend?

A.4 Topic number 4: Demographics of the U.S. Congress 2020

A.4.1 Text

The Democratic and Republican parties in the USA consist of many representatives with different religions, races and genders. However, the Democratic representatives seem to represent the general US population better. For example, the general US American population consists of even a little bit more women than men. While roughly 37% of the Democratic representatives are women, only about 10% of Republican representatives are. Concerning different races, a little bit over 60% of the population are White, which is about the same percentage as White Democratic representatives. However, more than 90% of the Republican representatives are White, so they are hugely over represented in this party. The second and third biggest groups in the population are Latinos/Hispanics (about 18%) and Blacks (about 13%). While Latinos/Hispanics are a little bit under represented in the Democratic representatives (roughly 15%), Blacks are a bit over represented with about 20%. In the Republican representatives, however, both are hugely under represented with about 4% of Latinos/Hispanics and only about 1% Blacks. People from Asia and the Pacific Islands make about 6% of the U.S. population and are represented with a roughly equal percentage in Democratic representatives, but again underrepresented in Republican representatives with only about 1%. Native Americans (roughly 1.5% of the population) are underrepresented in both parties' representatives. Concerning the religion, roughly 43% of Americans are Protestants, they are represented with a roughly equal percentage in Democratic representatives. However, a huge majority of about 70% of Republican representatives are Protestants. About a quarter of the population is unaffiliated religion-wise, which accounts for the second biggest group. This percentage is not represented by any of the two parties as in both the percentage of religious unaffiliated representatives is very small (under 1%) or non existant. The third biggest religious group in the population are Catholics with about 20%, but they are over represented in both parties with a bit over 30% for Democratic and roughly 25% of Republican representatives. While the Jewish religion is over represented and the third largest religion in Democratic representatives with about 10%, though they only account for 2% of the population, in the Republican representatives Mormons are overrepresented with about 4% even though they also only account for roughly 2% of the population.

A.4.2 Visualization

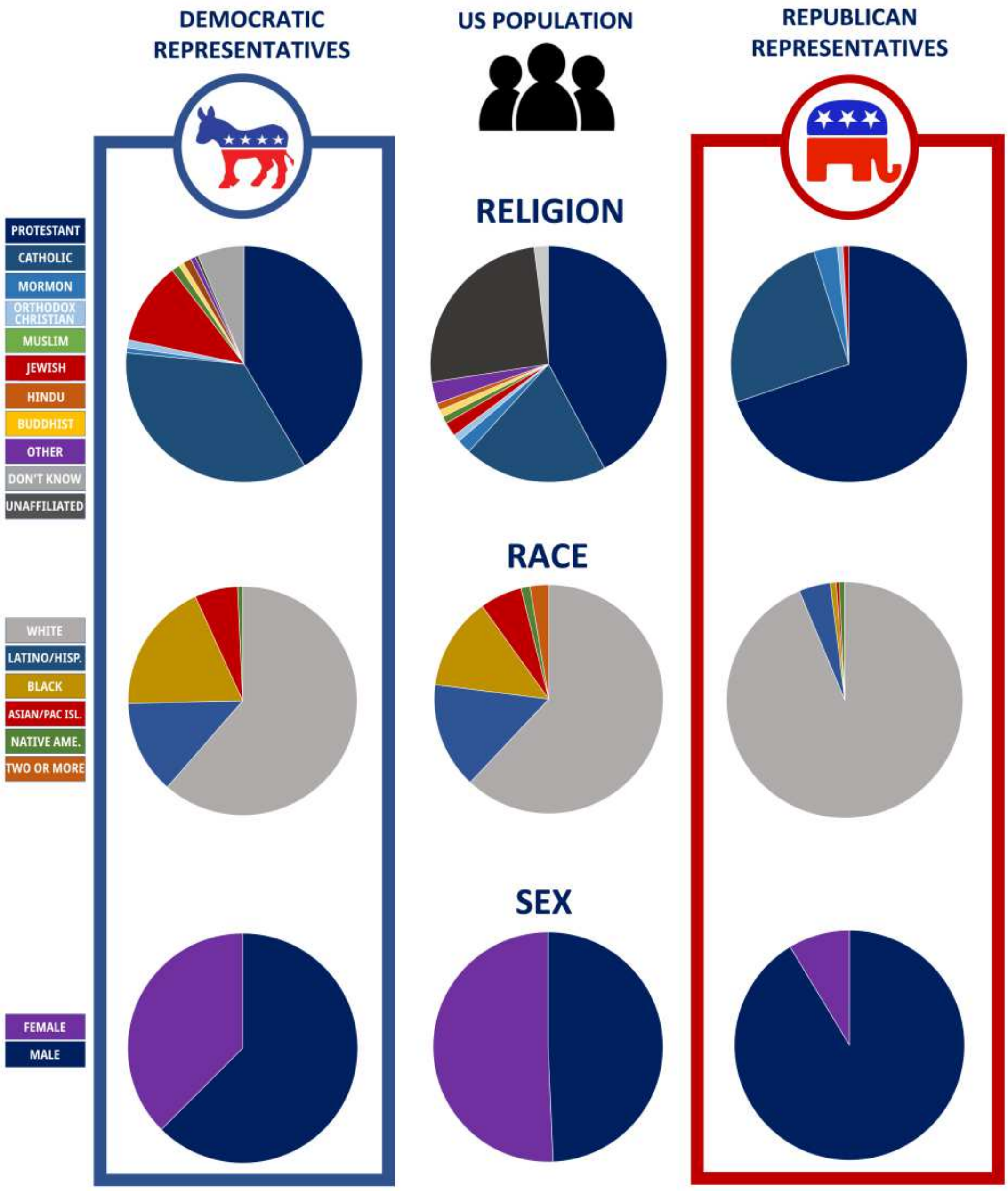


Figure A.4: Visualization for topic number 4 [V4].

A.4.3 Questions

1) Which race/s is/are over represented in the Democratic representatives in comparison to the population?

2) Approximately what percentage of Republican representatives are Catholics?
3) What is the third largest religion in terms of Democratic representatives?

A.5 Topic number 5: Revenue streams of Disney 2020

A.5.1 Text

The Walt Disney Company has many revenue streams. 2005 Bob Iger became the CEO of the company. Back then the revenue was a bit more than 30 billion dollars, but it more than doubled and rose to 70 billion dollars in 2019. Of these 70 billion dollars roughly 37% were earned by parks, experiences and products. The second highest portion (roughly 35%) was generated by earnings from media networks. Studio entertainment contributed roughly 16% and roughly 13% of the revenues were earned by DTC (direct to customer) and international revenues. Between the years 2005 and 2008, the revenues rose from a bit over 30 billion dollars to roughly 38 billion dollars. However, in 2009 there was a small decrease in revenues, where they only reached roughly 35 billion dollars. After that, revenues rose again until they were roughly 55 billion dollars in 2016. In 2017 revenues dropped again a little bit to roughly 54 billion dollars. However, in 2018 they increased to roughly 60 billion dollars. The biggest jump in revenues in the years 2005 to 2019 was between 2018 and 2019 where they rose from roughly 60 billion dollars to 70 billion dollars.

A.5.2 Visualization

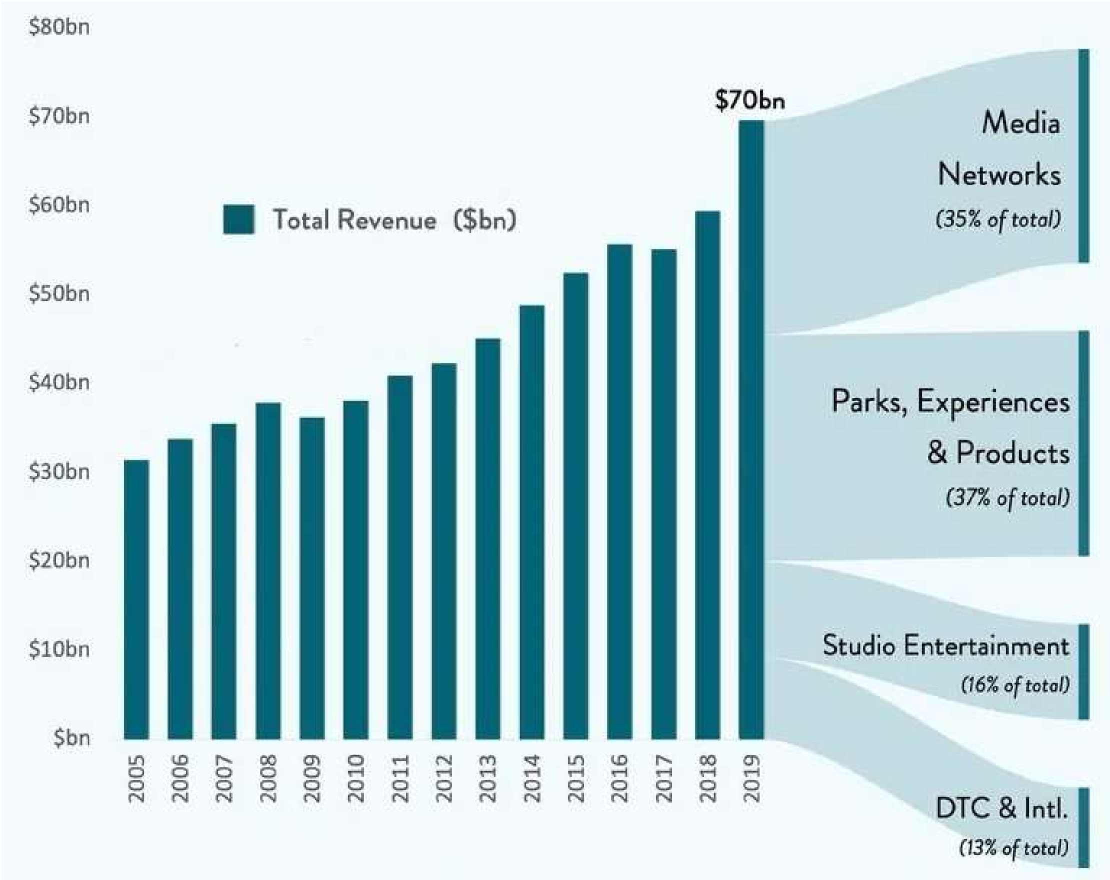


Figure A.5: Visualization for topic number 5 [V5].

A.5.3 Questions

1) In what year(s) did the revenue decrease from the previous year?
2) Which area created the most revenue in 2019?
3) What was the approximate amount of revenue in 2018?

A.6 Topic number 6: Elections in Austria

A.6.1 Text

An analysis of voting trends for Austria's Nationalratswahlen in 2013, 2017 and 2019 shows that in 2013 the biggest portion of voters were nonvoters. Roughly 50% of them stayed nonvoters in the next election in 2017, where there were nearly as many nonvoters as voters for SPÖ and FPÖ. Even a bit more than half of them stayed nonvoters in the election of 2019, where they were the second biggest portion of voters. However, the biggest part of nonvoters from 2013 that chose to vote in 2017 voted for the FPÖ, followed by ÖVP and SPÖ. In 2017 most voters chose to vote for the same party as in 2013. This doesn't account for the GRÜNE party, of which most voters went to other parties, especially to SPÖ, nonvoters, JETZT and ÖVP. For JETZT 2017 was the first year where they participated in the elections. However, they lost roughly 2/3 of voters in the next election 2019. Most of them went back to the GRÜNE party, which in general gained most of their voters in 2019 from the SPÖ and got more votes than in 2013 and 2017. While in 2013 the SPÖ got the most votes, they lost many voters to ÖVP, FPÖ and nonvoters and also some to JETZT and NEOS in 2017. Even though they still got roughly the same percentage of votes in 2017 compared to 2013, the ÖVP won the election in 2017, because they gained more votes from different parties. Apart from the voters from SPÖ, the ÖVP gained most from nonvoters, FPÖ and other parties. In 2019 the ÖVP still got most of the votes, of which many came from former FPÖ voters, nonvoters and former SPÖ voters. Compared to 2017, the SPÖ lost many voters to GRÜNE, nonvoters and ÖVP in 2019. While the FPÖ mostly gained votes in 2017 from nonvoters, other parties and SPÖ compared to 2013, they lost many voters in 2019. Most of them (roughly 30%) became nonvoters or chose to change to the ÖVP. While the NEOS lost some of their voters to the ÖVP in 2017 compared to 2013, they gained back more from the ÖVP in 2019, where roughly a third of NEOS voters were former ÖVP voters.

A.6.2 Visualization

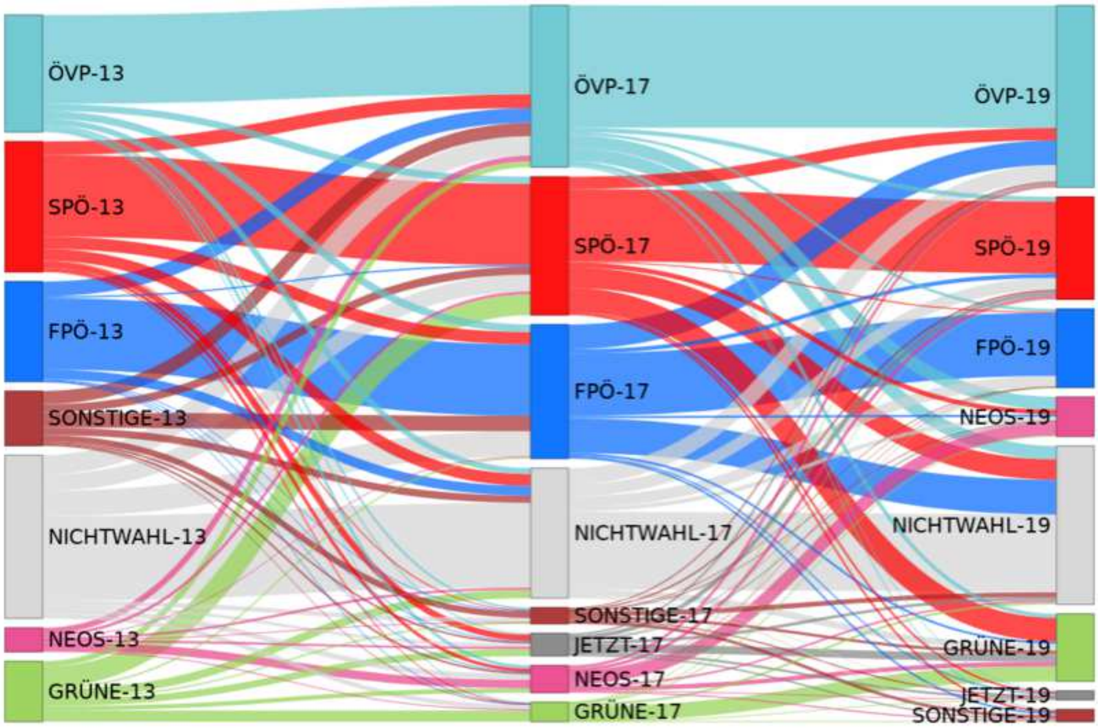


Figure A.6: Visualization for topic number 6 [V6].

A.6.3 Questions

1) Which party lost a large share of its voters to other parties in 2017 compared to 2013?
2) To which voter group did the FPÖ lose the most votes in 2019 compared to the previous election?
3) Which party got the most votes in 2017?

A.7 Topic number 7: Food waste in the UK

A.7.1 Text

All in all more than 9 million tonnes of food get discarded every year in the UK. More than half of the discarded food comes from homes. Thus, homes are by far the largest cause for food waste with 6.6 Million tonnes per year. A smaller portion of food waste is generated by food production with 1.5 million tonnes per year, followed by restaurants and takeaways with 1.1 million tonnes per year. Only a very small and neglectable percentage of food waste stems from retail. 54% or 5.1 tonnes of the waste food from all causes are incinerated. Some of this is used to generate electricity or is burnt to get fuel, but some is also released into the atmosphere. More than half of the incinerated food waste stems from private homes, followed by food production and restaurants and takeaways. A small percentage stems from retail. 26% or 2.5 million tonnes of food get dumped, of which 1 million tonnes land in landfill and the bigger portion or 1.5 million tonnes is dumped into the sewer systems. The problem with this is that the rotting food in landfills generates methane, which is a greenhouse gas. Also, dumping food into the sewer system can lead to problems like blockages and damaging of the pipes, which can cause high costs of repair. More than 2 million tonnes of the food that gets dumped stems from private homes, a small percentage from restaurants and takeaways and a very small percentage from food production. Only 20% or 1.9 million tonnes of food get recycled. This food is then redistributed to food banks and vulnerable people. Most of it comes from private homes that only recycle roughly 15% of their waste, followed by food production, which recycles nearly one third of its food waste, and retail, which recycles half of its waste. A small portion is from restaurants and takeaways.

A.7.2 Visualization

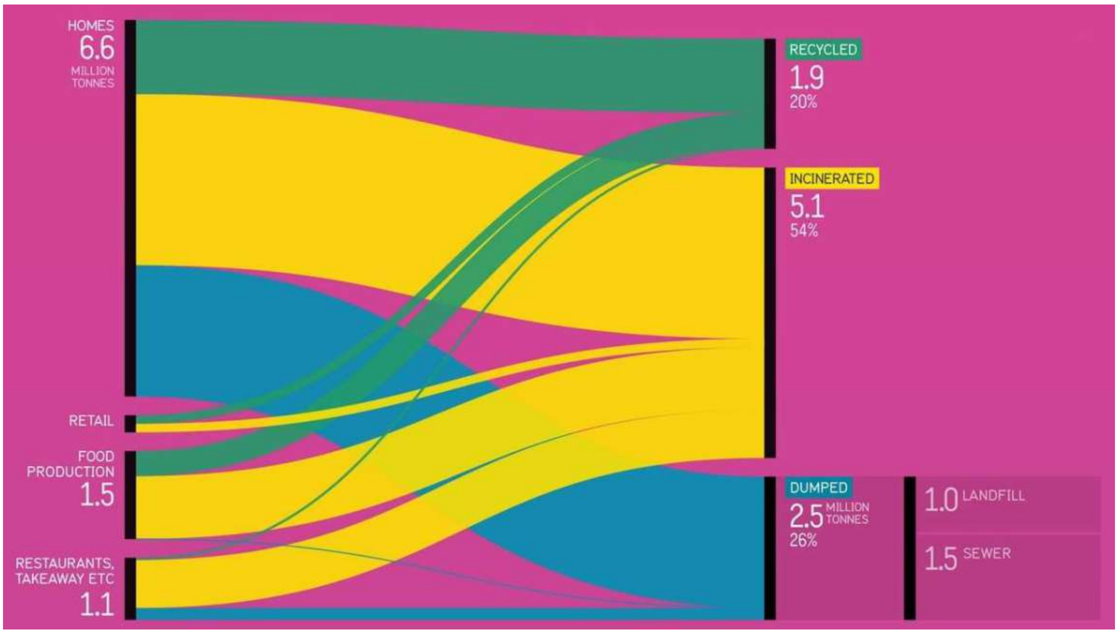


Figure A.7: Visualization for topic number 7 [V7].

A.7.3 Questions

1) From which source is the second most food recycled?
2) How many tonnes of food end up in the sewer?
3) What is the percentage difference between the amount of food that is incinerated and food that is recycled?

A.8 Topic number 8: Immigration to the U.S. 1820 - 2015

A.8.1 Text

Between the years 1820 and 2015 there were four immigration waves to the US, and there were more immigrants in each consecutive wave. The first wave was around the 1850s to the 1870s with roughly 2.8 million immigrants at the peak in around 1859. Back then, most immigrants were from Ireland (roughly 1 million on top of the wave) and Germany (roughly 980 thousand on top of the wave), followed by the United Kingdom (roughly 450 thousand on top of the wave). In the second wave that lasted from around the late 1870s to late 1890s, again, most people came from Germany with roughly 1.5 million people at peak, followed by the United Kingdom with roughly 800 thousand on top of the wave and Ireland with roughly 670 thousand on top of the wave. In total there were roughly 5.2 million immigrants at the peak around 1889. After the first two waves, the countries of origin of most of the immigrants that came to the US changed. The third wave was around 1900 to late 1930s. With roughly 2 million people at peak times, most came from Austria-Hungary, followed by Italy with 1.9 million at peak times and Russia with 1.5 million at peak times. In total, there were 8.2 million immigrants at the peak in 1909. During the years of the Second World War there was the least immigration to the US. Afterwards, the fourth wave started. It had its peak around 2008 where over 10 million people came to the US and again, the countries of origin were different than in the first three waves. In the fourth wave most people came from Mexico with roughly 2.8 million at peak times, Asian countries other than China, Russia, Philippines or India with roughly 1.7 million at peak times and the Caribbean with roughly 1 million at peak times.

A.8.2 Visualization

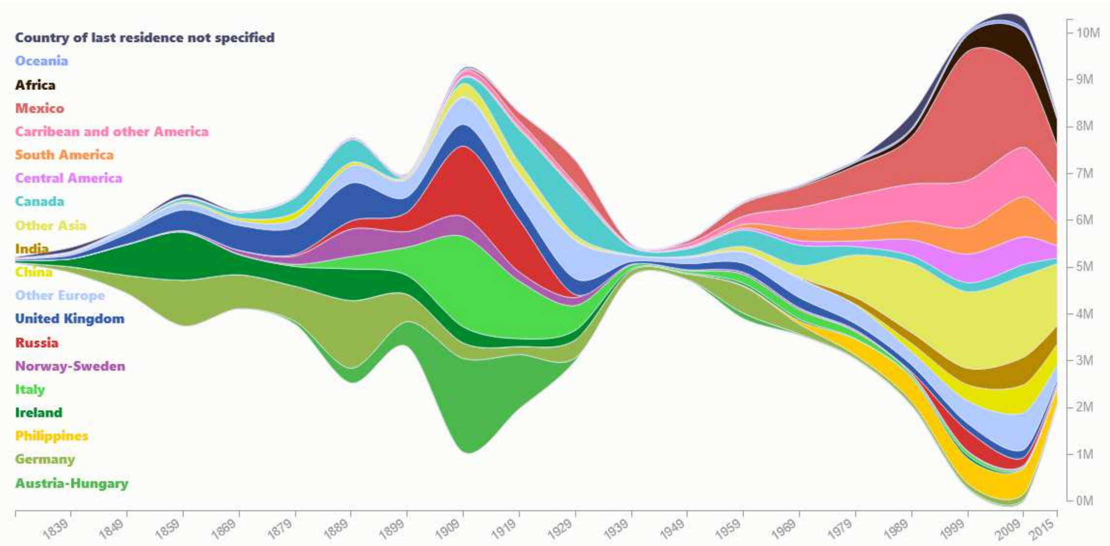


Figure A.8: Visualization for topic number 8 [V8].

A.8.3 Questions

1) Approximately how many people immigrated to the U.S. in total in 2008?
2) Estimated, how many people came from Austria-Hungary during the third wave?
3) During which period between 1850 and 1990 was there the least immigration?

A.9 Topic number 9: Global migration 2010 - 2015

A.9.1 Text

When looking at the global migration counts from 2010 to 2015, it gets clear that most of the time, people migrated within the same continent, and not from one continent to another. Exceptions were Oceania, where most immigrants (roughly 0.3 million of roughly 1.5 million migrants total) were from East Asia, as well as North America, where most immigrants (around 2 million of about 8 million migrants total) were from Latin America. Africa, South Asia, East Asia and Latin America had more emigration than immigration. The continent with the highest emigration counts was South Asia with nearly 8 million people (while its total migration was roughly 9 million people), of which about 3 million people migrated to West Asia, about 1.5 million to North America and roughly a little less than 1.5 million stayed in South Asia. A little less than a million of South Asians migrated to Europe and smaller portions to East Asia and Oceania. The continent with the highest percentage of immigration compared to its total migration was North America with roughly 7.3 million immigrants of a total of about 8 million migrants. Roughly 2 million people immigrated there from Latin America, about 1.5 million from East Asia and another 1.5 million from South Asia. Less than 1 million people came from Africa and even less from Europe and other continents. The continent with the highest migration in total was West Asia with nearly 15 million migrants in total, of which about 6 million were emigrants and 9 million immigrants. This was also the highest absolute number of immigrants for a continent. Of these 6 and 9 million people, 4.5 moved within West Asia. Only under half a million people went to Europe and even less to other continents. Roughly 3 million people migrated to West Asia from South Asia. This was the biggest amount of people moving from a continent to a different continent. Europe had a total migration count of about 11 million people, of which roughly 4 million were emigrants and 7 million immigrants. Roughly 2 million of both were people moving to a different place within Europe. About 1.7 million people migrated to Europe from Africa, and about 1 million from South Asia. Less than 0.5 million people came from East Europe/Central Asia, East Asia and West Asia, and even less from the other continents. About half a million people moved from Europe to East Europe/Central Asia and a little less to North America.

A.9.2 Visualization

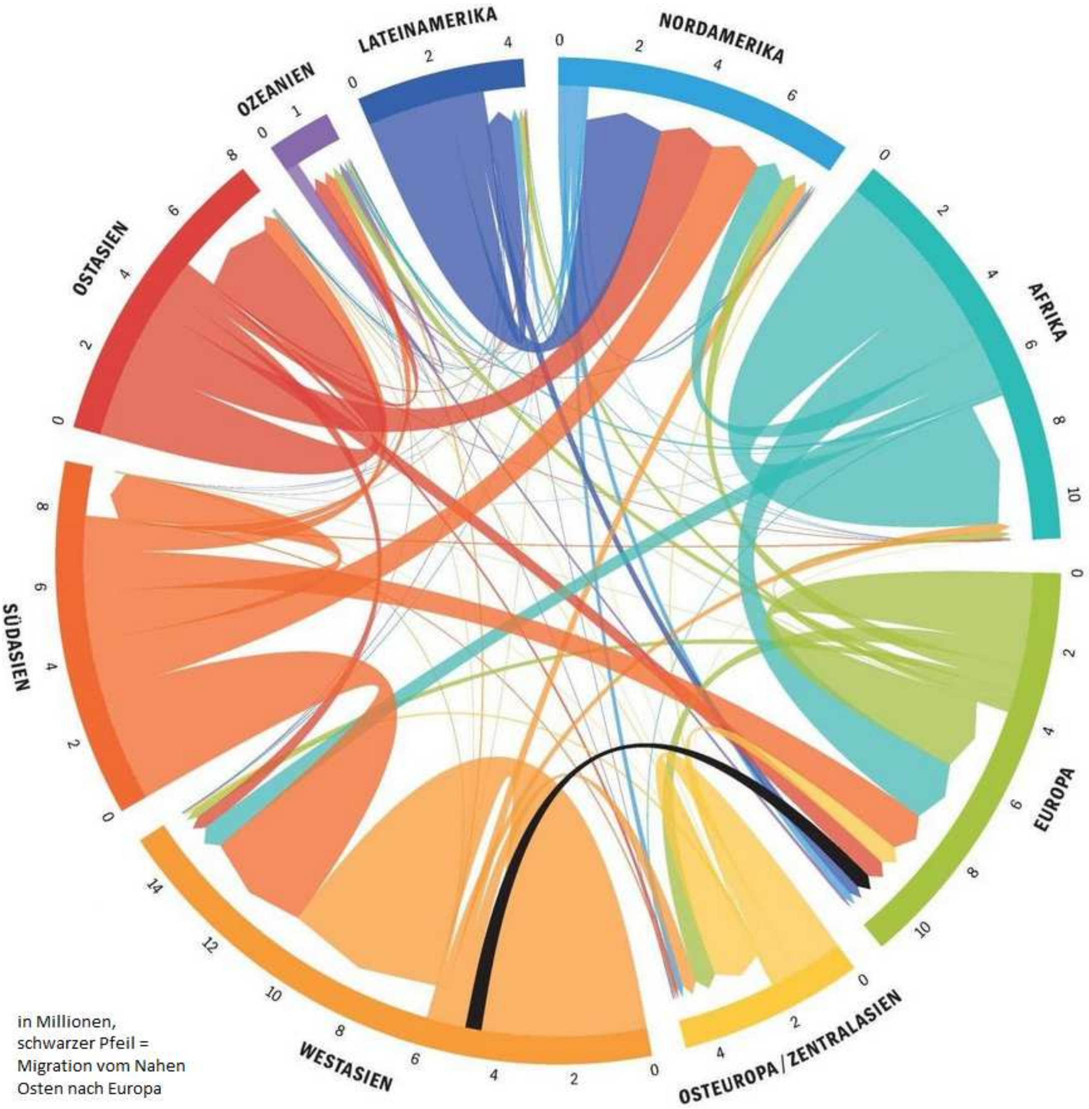


Figure A.9: Visualization for topic number 9 [V9].

A.9.3 Questions

1) Estimated how many immigrants came to North America from South Asia?
2) Which continent had the highest absolute number of immigrants?
3) Which continents recorded more emigration than immigration?

A.10 Topic number 10: Austria's mountains

A.10.1 Text

According to the "Österreichischer Alpenverein", in total, there are 3,225 mountains in Austria, which are higher than 999 meters. Some of them share common phrases in their names like "Groß...", "...berg" or "...kopf". However, there are also mountains with more unique names, like the Piz Buin, which is the highest mountain in Vorarlberg. The most common phrase in names is "...spitz(e)", which is a part of the names of 561 mountains. The mountains with this name are on average 2,630 meters high. The second and third most common phrase in mountain names are "...kogel" and "...berg", shared by 502 and 340 mountains respectively. The "...kogel"s are on average 2,056 meters high and the "...berg"e on average 1,599 meters. The phrase that is shared by the highest mountains on average is "...blick" with an average height of 2,935 meters. However, there are only four mountain names containing this phrase, which makes it to the phrase with the smallest number of mountains sharing it. The name parts of the second and third highest mountains on average are "Vordere/r..." with 2,741 meters on average and "Hintere/r..." with 2.689 meters on average. They are shared by 18 mountains, which is the second smallest number of mountains sharing a name part and 34 mountains, which is the fourth smallest number of mountains sharing their name. So it can be said, that the most common names are not shared by the on average highest mountains in Austria. Austria's highest mountain, the Großglockner has the same name part "Groß(e/r)..." as 121 mountains, which is only the eighth most common phrase out of 18. Concerning the average height, the mountains named "Groß(e/r)..." are also only on the ninth place out of 18 with 2,324 meters on average.

A.10.2 Visualization

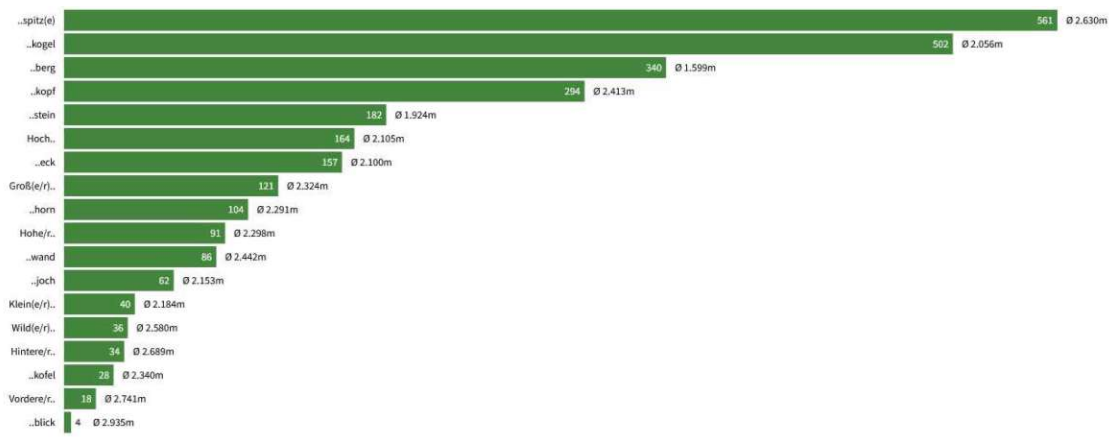


Figure A.10: Visualization for topic number 10 [V10].

A.10.3 Questions

1) The mountains with which name component are the second highest on average?
2) What is the average height of mountains with the name component "Groß(e/r)"?
3) How many mountains have "...berg" in their name?

A.11 Topic number 11: Austria's population

A.11.1 Text

In the year 2021 Austria had a population of 8,960,751 people. 84,141 of these were children between 0 and 1 years and a little bit less children (84,039) between 1 and 2 years. From all the people in Austria, who share the same age and were between 0 and 74 years old, the groups 0-1 and 1-2 years contained the fewest people. In general it can be said that there are less and less younger and more older people in Austria. The birth rates have mostly decreased since 1969 and more or less stagnated over the last 20 years. Looking at the years between 2011 and 2021 the year in which most people were born was 2017 with 89,855. In the last ten years a little bit more boys than girls were born. In general, in the age group 0 to 45 there are more men than women in Austria. From age 46, there is mostly an excess of women with exceptions from ages 53 to 55 where there are a little bit more men. The excess of women is especially big from ages 67 and upwards, where for every age, there are at least 5,000 more women than men. Roughly from ages 89 and up, there are even at least twice as much women than men. In 2021 Austria's population contained only 1,301 people who were born in 1921. Most people in Austria were born in the years 1964 to 1969 with over 140,000 people each year. The amount of people with a certain age naturally decreases from years 80 and up. Apart from that, people that were born in 1946 and were 75 years old were the smallest age group in Austria in 2021 with 62,739 persons.

A.11.2 Visualization

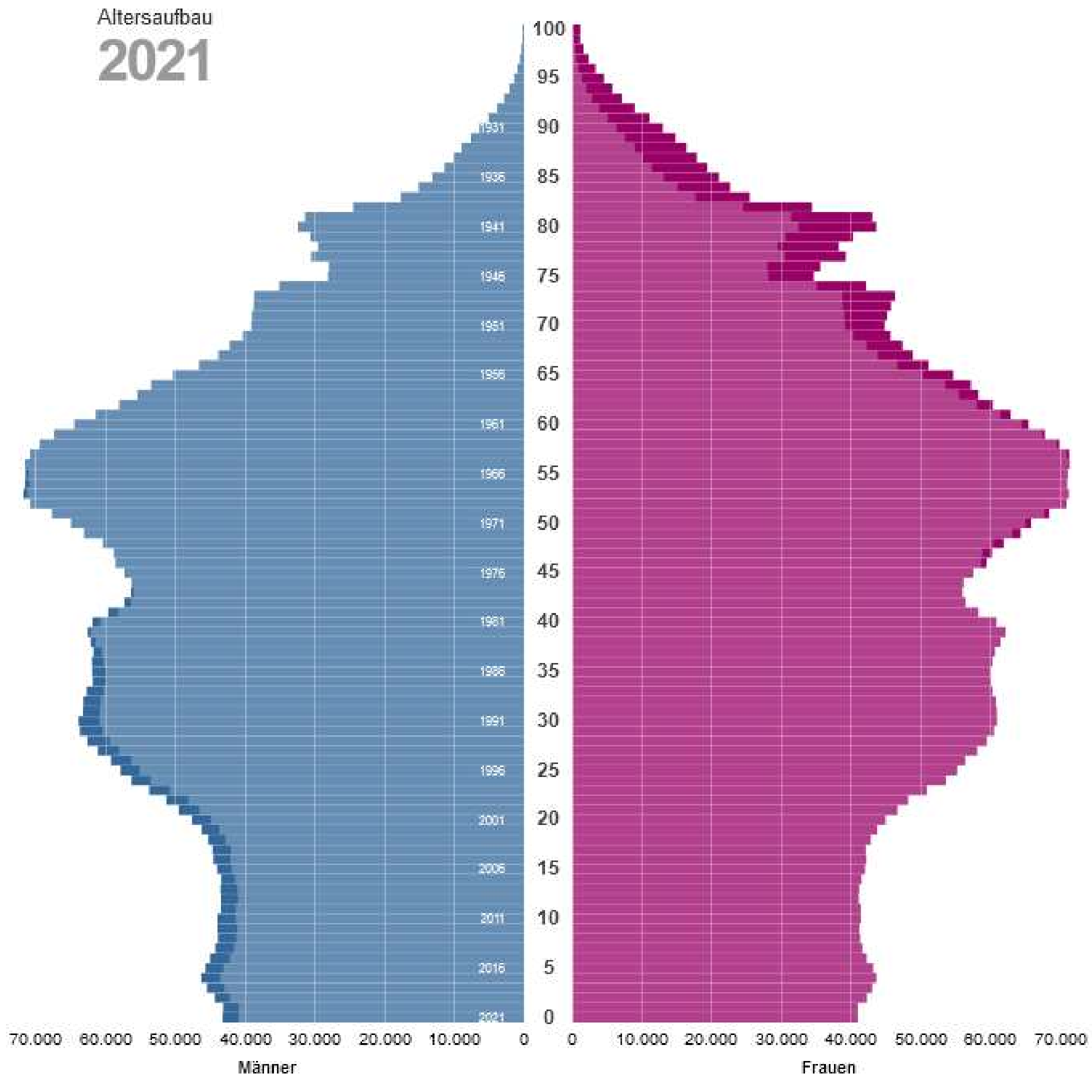


Figure A.11: Visualization for topic number 11 [V11].

A.11.3 Questions

1) Of those between 0 and 80 years old, what year was the smallest age group from?
2) From what age on, roughly, are there twice as many women as men?
3) Approximately how many people are from 1967?

A.12 Topic number 12: Life expectancy per continent 2016

A.12.1 Text

The life expectancy is very different from continent to continent. There can be also differences in life expectancy between countries from the same continent, which are more or less developed than others, leading to multiple peaks in life expectancy curves per continent. In 2016 people in Europe had the highest life expectancy with roughly 82 years in most European countries and in many others at least over 75 years. Africa had the least life expectancy, which was only roughly 62 years in most countries. However, there were also some African countries with a life expectancy of nearly 75 years, while others only had a life expectancy of 50-55 years. Most countries in Asia had a life expectancy of roughly 76 years, directly followed by many Asian countries with a life expectancy of 70-75 years, but there were also some Asian cuntries with life expectancies over 83 years. In Oceania life expectancy was roughly 71 years in most countries, but roughly 83 years in others. Most countries in North and South America had life expectancies of roughly 75 years. However, in North America, there was a higher amount of countries having a higher life expectancy than 75 years than in South America. While in South America, there were more countries having a lower life expectancy than 75 years than in North America. Though, in North America there were more countries with a life expectancy in the range between 60 and 65 years, which is quite low.

A.12.2 Visualization

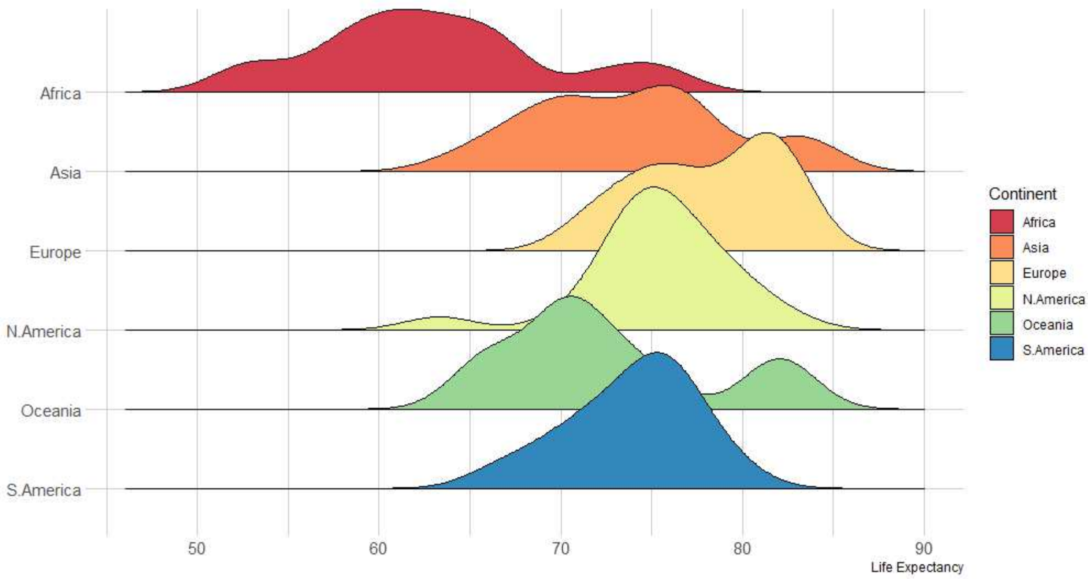


Figure A.12: Visualization for topic number 12 [V12].

A.12.3 Questions

1) What is the approximate life expectancy in South America?
2) Which continent has the highest life expectancy?
3) Approximately, what is the difference in life expectancy between most african countries and most european countries?

A.13 Topic number 13: Gender pay gap in the U.S.

A.13.1 Text

Nowadays, in most jobs, men still earn more than women, even though they do the same amount of work. Statistics showed the pay gaps of 26 jobs from three different job categories in the US. The categories were "Admin & Organization", "Care & Education" and "Law & Justice". It can be seen that the pay gap was especially big for doctors and surgeons, where men earned the most with roughly 105k dollars per year, while women only earned about 65k dollars, leading to a pay gap of about 40k dollars. Legal occupations had the same pay gap, with men earning about 92k dollars and women with about 52k dollars. HR managers had the third biggest pay gap, where men had a yearly salary of about 95k dollars, while women only earned about 68k dollars, resulting in a pay gap of about 27k dollars. The jobs with the smallest pay gaps in this statistic were traffic clerks, which only had a pay gap of about 2k dollars, while they earned around 30k dollars, office clerks with a pay gap of about 3k dollars, while they earned between 30k and 35k dollars and social workers with a pay gap of about 4k dollars while they earned about 45k dollars. Three of the four jobs in which men earned most also belonged to the top 3 jobs with the biggest pay gaps, which were mentioned before. The fourth one were lawyers, which was the second best paid job for men with roughly 100k dollars per year. Female lawyers earned roughly 83k dollars per year, which lead to it being the best paying job for women. The second best paying job for women was physical therapist with about 68k dollars per year, while men in the same job earned about 9k dollars more. Despite the third largest pay gap, HR manager was still the third best paying job for women, followed by doctor and surgeon which even had the highest pay gap. The three least paying jobs were the same for men and women, namely nursing, healthcare support and teacher assistant. They all had pay gaps of between roughly 3k and 5k dollars per year and salaries between about 24k and 26k for women and about 28k to 30k for men. In none of the 26 jobs women earned more than men.

A.13.2 Visualization

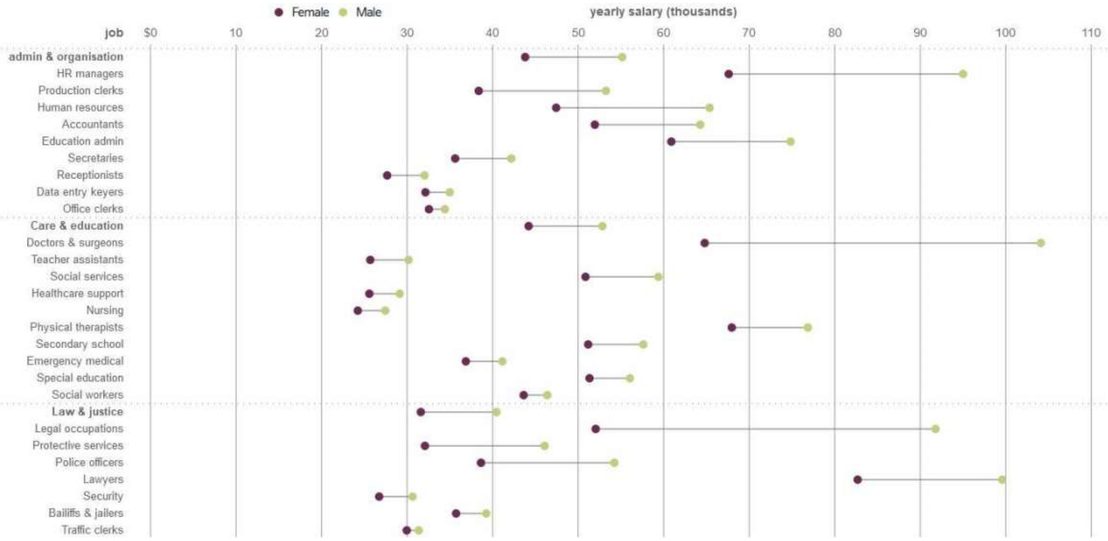


Figure A.13: Visualization for topic number 13 [V13].

A.13.3 Questions

1) In which profession did women earn the fourthmost?
2) Approximately what is the difference (in dollars) between the highest paying job for men and the highest paying job for women?
3) Which job had the second smallest pay gap?

A.14 Topic number 14: Places of residence and work in Vienna

A.14.1 Text

People in Vienna don't necessarily work in the same district as they live in. A statistic from 2018 showed that only in the 1st district more than 50% of people worked either at home (25%) or in the same district (27%). Only 21% of residents worked in a neighbouring district, 17% in an other district and 10% not in Vienna. However, the numbers for the remaining districts were very different. The districts with second and third most people working from home or in the same district were the 3rd and 9th with only 25% and 22% respectively. The districts in which the least residents worked from home or in the same district were the 20th with 13%, and the 15th and 17th with both 14%. The district where the least people worked from home was the 11th with only 4%, while 11% worked in the same district. In the 8th district, the least people worked in the same district (6%) and most (41%) in a different district than a neighbouring one. The 23rd had the highest number of people working outside of Vienna (16%), while in the other districts mostly around 10% of residents didn't work in Vienna. In the 22nd district, the fewest people, which were only 6%, worked in the neighbouring district and also the most people, or 62%, worked in a different district than a neighbouring one compared to all the other districts. Whereas the 4th district had the highest percentage of people working in a neighbouring district (34%) compared to the others. In general it can be said that in 2018 in a majority of Viennese districts most people worked in a different district other than a neighbouring one. The percentages were especially high in the districts 5, 14, 16, 18, 19, 20, 21, 22, where they were between 57% and 62%.

A.14.2 Visualization

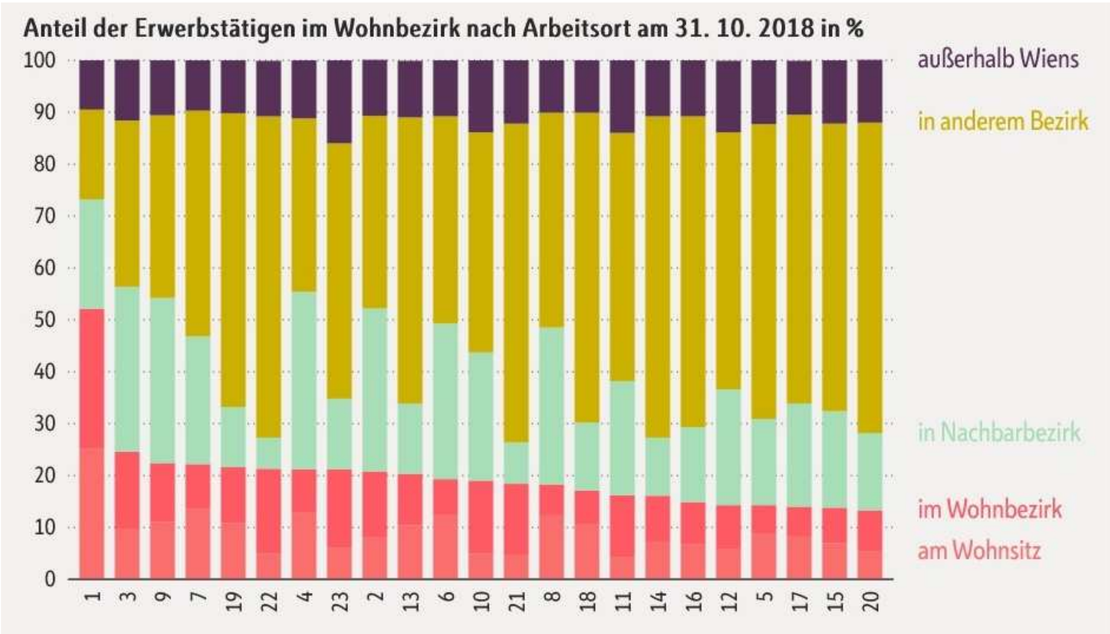


Figure A.14: Visualization for topic number 14 [V14].

A.14.3 Questions

1) In which district did most people work in a district other than their neighboring district (in anderem Bezirk)?
2) Which district had the fewest people working at home (am Wohnsitz)?
3) Approximately what percentage of people from the 23rd district worked outside of Vienna?

A.15 Topic number 15: Yearly inflation 2021

A.15.1 Text

Comparing the inflation across the USA and the EU in 2021 it gets clear that all in all, there was less inflation in the EU. However, the range within the EU was bigger, where some countries had an inflation of 3% and less, while others had 8% and more. In the USA the inflation was between 6% and 8% and more in all of the states. There, the inflation was especially high in many states in the west of the USA, where it was mostly 8% or more. An Exception was the west coast, where the inflation only reached 6%. Another part of the USA with an inflation of only about 6% was the northeast. Countries belonging to the EU having had the lowest inflation of 3% were for example France and Portugal. Some countries with 4% inflation were Austria, Italy, Greece, Denmark, Sweden or Finland. Ireland, the Czech Republic, Slovakia, Slovenia, Croatia and Cyprus had an inflation of 5%. Other countries of the EU had an equally high inflation as some parts of the USA. For example, Spain, Germany, Romania, the Netherlands and Luxembourg had 6%, while Belgium, Latvia, Poland and Bulgaria even had an inflation of 7%. However, the highest inflation rates in the EU were reached by Estonia, Lithuania and Hungary with 8% or more. In general, especially countries from East Europe had higher inflation rates than the others. Some other European Countries that don't belong to the EU but had lower inflation rates than the USA were Switzerland with 3% or less, Iceland and UK with 4% and Norway with 5%.

A.15.2 Visualization

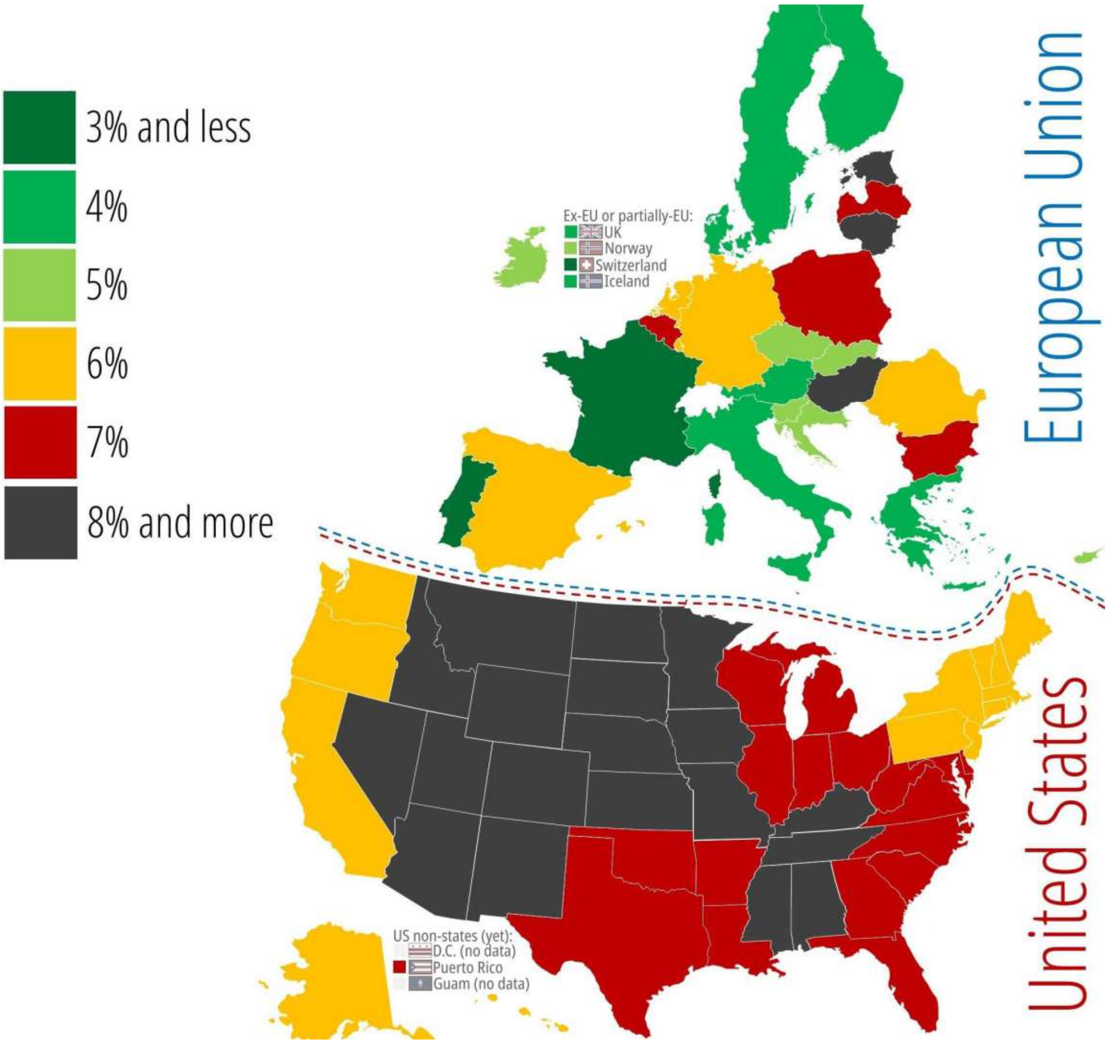


Figure A.15: Visualization for topic number 15 [V15].

A.15.3 Questions

1) Name two of the European countries where inflation was 3% or less.
2) Approximately what was the difference in percentage terms between the lowest inflation in Europe and the lowest in the U.S.?
3) Was inflation in Austria higher or lower than that in Norway?

List of Figures

List of Tables

Bibliography

[3M97] 3M. Polishing your Presentation. http://web.archive.org/web/20001102203936if_/http://3m.com:80/meetingnetwork/files/meetingguide_pres.pdf, 1997. Accessed: 2022-10-17.

[AAJ+21] Louis Allen, Jack Atkinson, Dinusha Jayasundara, Joan Cordiner, and Peyman Z Moghadam. Data visualization for Industry 4.0: A stepping-stone toward a digital future, bridging the gap between academia and industry. *Patterns*, 2(5):100266, 2021.

[Ans73] Francis J. Anscombe. Graphs in statistical analysis. *The American Statistician*, 27(1):17–21, 1973.

[AWR+18] Yuliya Ardasheva, Zhe Wang, Anna Karin Roo, Olusola O. Adesope, and Judith A. Morrison. Representation visuals' impacts on science interest and reading comprehension of adolescent English learners. *The Journal of Educational Research*, 111(5):631–643, 2018.

[BBG19] Katy Börner, Andreas Bueckle, and Michael Ginda. Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6):1857–1864, 2019.

[BBK+15] Michelle A. Borkin, Zoya Bylinskii, Nam Wook Kim, Constance May Bainbridge, Chelsea S. Yeh, Daniel Borkin, Hanspeter Pfister, and Aude Oliva. Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528, 2015.

[BDDG03] Patrick Baudisch, Doug DeCarlo, Andrew T. Duchowski, and Wilson S. Geisler. Focusing on the essential: considering attention in display design. *Communications of the ACM*, 46(3):60–66, 2003.

[Ben09] BenRG. Cone fundamentals with SRGB spectrum. https://commons.wikimedia.org/w/index.php?curid=7873848, 2009. Accessed: 2022-10-03.

[Bor13] Michelle A. Borkin. Different types of visualizations. https://seas.harvard.edu/sites/default/files/images/news/visualizations.jpg, 2013. Accessed: 2023-01-04.

[BR00] Robert Maribe Branch and Jean Ann Riordan. Time and the use of diagrams or texts, and study questions on learner comprehension. *Journal of Visual Literacy*, 20(2):197–218, 2000.

[BRBF14] Jeremy Boy, Ronald A. Rensink, Enrico Bertini, and Jean-Daniel Fekete. A principled way of assessing visualization literacy. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1963–1972, 2014.

[BSS+18] Michael Behrisch, Dirk Streeb, Florian Stoffel, Daniel Seebacher, Brian Matejek, Stefan Hagen Weber, Sebastian Mittelstädt, Hanspeter Pfister, and Daniel Keim. Commercial visual analytics systems–advances in the big data analytics field. *IEEE Transactions on Visualization and Computer Graphics*, 25(10):3011–3031, 2018.

[BVB+13] Michelle A. Borkin, Azalea A. Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2306–2315, 2013.

[BvL09] Inez E. Berends and Ernest C.D.M. van Lieshout. The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning and Instruction*, 19(4):345–353, 2009.

[Car09] Stuart Card. Information visualization. In *Human-Computer Interaction*, pages 199–234. CRC Press, 2009.

[CEH+08] Min Chen, David Ebert, Hans Hagen, Robert S Laramee, Robert Van Liere, Kwan-Liu Ma, William Ribarsky, Gerik Scheuermann, and Deborah Silver. Data, information, and knowledge in visualization. *IEEE Computer Graphics and Applications*, 29(1):12–19, 2008.

[CGW15] Wei Chen, Fangzhou Guo, and Fei-Yue Wang. A survey of traffic data visualization. *IEEE Transactions on Intelligent Transportation Systems*, 16(6):2970–2984, 2015.

[Chi19] Nayomi Chibana. Do infographics increase reading comprehension and recall? `https://visme.co/blog/infographics-vs-text/`, 2019. Accessed: 2022-02-02.

[CM84] William S. Cleveland and Robert McGill. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *Journal of the American Statistical Association*, 79(387):531–554, 1984.

[CM97] Stuart Card and Jock Mackinlay. The structure of the information visualization design space. In *Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium*, pages 92–99. IEEE, 1997.

[CMS99] Stuart Card, Jock Mackinlay, and Ben Shneiderman. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann, 1999.

[Dah06] Markus Dahm. *Grundlagen der Mensch-Computer-Interaktion*. Pearson Studium München, 2006.

[DH00] Mary Dyson and Mark Haselgrove. The effects of reading speed and reading patterns on the understanding of text read from screen. *Journal of Research in Reading*, 23(2):210–223, 2000.

[DIA] DIAGRAM Center Image Description Guidelines. `http://diagramcenter.org/table-of-contents-2.html`. Accessed: 2022-10-12.

[DML14] Maxime Dumas, Michael J. McGuffin, and Victoria L. Lemieux. Finance-vis.net - A visual survey of financial data visualizations. In *Poster Abstracts of IEEE Conference on Visualization*, volume 2, page 8, 2014.

[Dun19] Matthew Dunn. Research: Is a picture worth 1,000 words or 60,000 words in marketing? `https://www.emailaudience.com/research-picture-worth-1000-words-marketing/`, 2019. Accessed: 2022-02-02.

[Ell18] Geoffrey Ellis. *Cognitive Biases in Visualizations*. Springer, 2018.

[Ent18] Entrepreneur Middle East Staff. Infographic: Five Scientific Reasons People Are Wired To Respond To Visual Marketing. `https://www.entrepreneur.com/en-ae/marketing/infographic-five-scientific-reasons-people-are-wired-to/312551`, 2018. Accessed: 2022-10-17.

[Erf12] Seyyed Mahdi Erfani. Pictures speak louder than words in ESP, too!. *English Language Teaching*, 5(8):164–169, 2012.

[Fec60] Gustav Theodor Fechner. *Elemente der Psychophysik, Volume 2*. Breitkopf u. Härtel, 1860.

[Few14] Stephen Few. Data visualization for human perception. *The Encyclopedia of Human-Computer Interaction, 2nd Ed.*, 2014.

[Fib07] Fibonacci. Müller-Lyer illusion. `https://commons.wikimedia.org/w/index.php?curid=1792612`, 2007. Accessed: 2022-10-03.

[Fır17] Mehmet Fırat. How real and model visuals affect the test performance of elementary students. *Computers in Human Behavior*, 71:258–265, 2017.

[FJL22a] Elif E. Firat, Alark Joshi, and Robert S. Laramee. Interactive visualization literacy: The state-of-the-art. *Information Visualization*, 21(3):285–310, 2022.

[FJL22b] Elif E. Firat, Alark Joshi, and Robert S. Laramee. VisLitE: Visualization Literacy and Evaluation. *IEEE Computer Graphics and Applications*, 42(3):99–107, 2022.

[Fot20] Fotis, Spiros Jr. The Power of Data Visualization. `https://aegisresearch.eu/the-power-of-data-visualization/`, 2020. Accessed: 2022-10-17.

[FW21] Michael Friendly and Howard Wainer. *A History of Data Visualization and Graphic Communication*. Harvard University Press, 2021.

[FWSN08] Jean-Daniel Fekete, Jarke J. van Wijk, John T. Stasko, and Chris North. The value of information visualization. In *Information Visualization*, pages 1–18. Springer, 2008.

[GMB16] Sam Grainger, Feng Mao, and Wouter Buytaert. Environmental data visualisation for non-scientific contexts: Literature review and design framework. *Environmental Modelling & Software*, 85:299–318, 2016.

[GZL+14] Qihong Gan, Min Zhu, Mingzhao Li, Ting Liang, Yu Cao, and Baoyao Zhou. Document visualization: An overview of current research. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(1):19–36, 2014.

[GZWM20] Daibao Guo, Shuai Zhang, Katherine Landau Wright, and Erin M. McTigue. Do you get the picture? A meta-analysis of the effect of graphics on reading comprehension. *American Educational Research Association Open*, 6(1):1–20, 2020.

[HDF+06] Olaf Hauk, Matthew H. Davis, M. Ford, Friedemann Pulvermüller, and William D. Marslen-Wilson. The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage*, 30(4):1383–1400, 2006.

[HE11] Christopher Healey and James Enns. Attention and visual memory in visualization and computer graphics. *IEEE Transactions on Visualization and Computer Graphics*, 18(7):1170–1188, 2011.

[HOWBN17] Mona Holmqvist Olander, Eva Wennås Brante, and Marcus Nyström. The effect of illustration on improving text comprehension in dyslexic adults. *Dyslexia*, 23(1):42–65, 2017.

[IKT+13] Kenji Ikeda, Shinji Kitagami, Tomoyo Takahashi, Yosuke Hattori, and Yuichi Ito. Neuroscientific information bias in metacomprehension: The effect of brain images on metacomprehension judgment of neuroscience research. *Psychonomic Bulletin & Review*, 20(6):1357–1363, 2013.

[ine] inetSoft. Executive Sales Summary Dashboard. https://www.inetsoft.com/evaluate/bi_visualization_gallery/dashboard.jsp?dbIdx=8. Accessed: 2023-02-06.

[JK17] Yu-Cin Jian and Hwa-Wei Ko. Influences of text difficulty and reading ability on learning illustrated science texts for children: An eye movement study. *Computers & Education*, 113:263–279, 2017.

[JMM+05] Chris Johnson, Robert Moorhead, Tamara Munzner, Hanspeter Pfister, Penny Rheingans, and Terry S. Yoo. NIH-NSF visualization research challenges report. Institute of Electrical and Electronics Engineers, 2005.

[KPHH12] Sean Kandel, Andreas Paepcke, Joseph M. Hellerstein, and Jeffrey Heer. Enterprise data analysis and visualization: An interview study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2917–2926, 2012.

[Lev12] Alan Levine. The 60,000 times question remains unanswered. https://cogdogblog.com/2012/07/60000-times-question/, 2012. Accessed: 2022-02-02.

[LKK16] Sukwon Lee, Sung-Hee Kim, and Bum Chul Kwon. VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):551–560, 2016.

[LL82] W. Howard Levie and Richard Lentz. Effects of text illustrations: A review of research. *Educational Communication and Technology*, 30(4):195–232, 1982.

[LSM13] Alwine Lenzner, Wolfgang Schnotz, and Andreas Müller. The role of decorative pictures in learning. *Instructional Science*, 41(5):811–831, 2013.

[mas15] MASSVIS dataset. http://massvis.mit.edu/, 2015. Accessed: 2022-04-11.

[Maz09] Riccardo Mazza. *Introduction to Information Visualization*. Springer Science & Business Media, 2009.

[MHD+17] Laura E. Matzen, Michael J. Haass, Kristin M. Divis, Zhiyuan Wang, and Andrew T. Wilson. Data visualization saliency model: A tool for evaluating abstract data visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):563–573, 2017.

[MSLP07] Matthew T. McCrudden, Gregory Schraw, Stephen Lehman, and Anne Poliquin. The effect of causal diagrams on text learning. *Contemporary Educational Psychology*, 32(3):367–388, 2007.

[Mun14] Tamara Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[MWW+22] Malissa Maria Mahmud, Shiau Foong Wong, Shiau San Wong, Othman Ismail, and Chandra Reka Ramachandiran. Data Visualization Literacy: Multimodal skills in the Age of Internet of Things (IoT). In *Proceedings of the 7th International Conference on Information and Education Innovations*, pages 47–53, 2022.

[NWC+21] Carolina Nobre, Dylan Wootton, Zach Cutler, Lane Harrison, Hanspeter Pfister, and Alexander Lex. revisit: Looking under the hood of interactive visualization studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–13, 2021.

[NYS09] Shahrul Azman Noah, Suraya Yaakob, and Suzana Shahar. Application of information visualization techniques in representing patients' temporal personal history data. In *International Visual Informatics Conference*, pages 168–179. Springer, 2009.

[Pla86] William Playfair. *The Commercial and Political Atlas: Representing, by Means of Stained Copper-Plate Charts, the Progress of the Commerce, Revenues, Expenditure and Debts of England During the Whole of the Eighteenth Century*. T. Burton, 1786.

[PP09] Yi-Chun Pan and Yi-Ching Pan. The effects of pictures on the reading comprehension of low-proficiency taiwanese english foreign language college students: An action research study. *VNU Journal of Foreign Studies*, 25(3), 2009.

[PR19] Tatiana E. Petrova and Elena I. Riekhakaynen. Processing of verbal and non-verbal patterns: an eye-tracking study of russian. In *Proceedings of the Third International Congress on Information and Communication Technology*, pages 269–276. Springer, 2019.

[PVF13] Charles Perin, Romain Vuillemot, and Jean-Daniel Fekete. Soccerstories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013.

[PWHM14] Mary C. Potter, Brad Wyble, Carl Erick Hagmann, and Emily S. McCourt. Detecting meaning in RSVP at 13 ms per picture. *Attention, Perception, & Psychophysics*, 76(2):270–279, 2014.

[PWV+18] Patrik Polatsek, Manuela Waldner, Ivan Viola, Peter Kapec, and Wanda Benesova. Exploring visual attention and saliency modeling for task-based visual analysis. *Computers & Graphics*, 72:26–38, 2018.

[Rhc07] Rhcastilhos and Jmarchn. Schematic diagram of the human eye with english annotations. `https://commons.wikimedia.org/w/index.php?curid=1597930`, 2007. Accessed: 2022-10-03.

[RHM88] David Reinking, David A. Hayes, and John E. McEneaney. Good and poor readers' use of explicitly cued graphic aids. *Journal of Reading Behavior*, 20(3):229–247, 1988.

[RNC15] Kathryn L. Roberts, Rebecca R. Norman, and Jaime Cocco. Relationship between graphical device comprehension and overall text comprehension for third-grade children. *Reading Psychology*, 36(5):389–420, 2015.

[RTdOT06] Jose F. Rodrigues, Agma J.M. Traina, Maria Cristina Ferreira de Oliveira, and Caetano Traina. Reviewing data visualization: An analytical taxonomical study. In *Tenth International Conference on Information Visualisation (IV'06)*, pages 713–720. IEEE, 2006.

[SD10] Michael J. Serra and John Dunlosky. Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory*, 18(7):698–711, 2010.

[Shn03] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *The Craft of Information Visualization*, pages 364–371. Elsevier, 2003.

[Sib12] Sibley, Amanda. 19 Reasons You Should Include Visual Content in Your Marketing [Data]. `https://blog.hubspot.com/blog/tabid/6307/bid/33423/19-reasons-you-should-include-visual-content-in-your-marketing-data.aspx`, 2012. Accessed: 2022-10-17.

[SKND22] Marc Satkowski, Franziska Kessler, Susanne Narciss, and Raimund Dachselt. Who benefits from Visualization Adaptations? Towards a better Understanding of the Influence of Visualization Literacy. *arXiv preprint arXiv:2207.05401*, 2022.

[SLU+14] Wolfgang Schnotz, Ulrich Ludewig, Mark Ullrich, Holger Horz, Nele McElvany, and Jürgen Baumert. Strategy shifts during learning from texts and pictures. *Journal of Educational Psychology*, 106(4):974, 2014.

[Spi22] Spillane, James. Visual Marketing: A Picture's Worth 60,000 Words. `https://www.business2community.com/digital-marketing/visual-marketing-pictures-worth-60000-words-01126256`, 2022. Accessed: 2022-10-17.

[Ste57] Stanley S. Stevens. On the psychophysical law. *Psychological Review*, 64(3):153, 1957.

[Ste01] Robert A. Stebbins. *Exploratory research in the social sciences, Volume 48*. Sage, 2001.

[TCC22] Marta Terrado, Luz Calvo, and Isadora Christel. Towards more effective visualisations in climate services: Good practices and recommendations. *Climatic Change*, 172(1):1–26, 2022.

[TFM96] Simon Thorpe, Denis Fize, and Catherine Marlot. Speed of processing in the human visual system. *Nature*, 381(6582):520–522, 1996.

[Tho16] Clive Thompson. The surprising history of the infographic. `https://www.smithsonianmag.com/history/surprising-history-infographic-180959563/`, 2016. Accessed: 2022-01-18.

[TM04] Melanie Tory and Torsten Möller. Rethinking visualization: A high-level taxonomy. In *IEEE Symposium on Information Visualization*, pages 151–158. IEEE, 2004.

[Tuf85] Edward R. Tufte. The visual display of quantitative information. *The Journal for Healthcare Quality (JHQ)*, 7(3):15, 1985.

[V1] Visualization number 1. `https://www.economist.com/graphic-detail/2012/08/20/puffed-out`. Accessed: 2023-02-06.

[V10] Visualization number 10. `https://www.drawingdata.net/berge/`. Accessed: 2022-01-20.

[V11] Visualization number 11. `https://www.statistik.at/atlas/bev_prognose/`. Accessed: 2023-02-06.

[V12] Visualization number 12. `https://twitter.com/Juanma_MN/status/1143972460913709056/photo/1`. Accessed: 2023-02-06.

[V13] Visualization number 13. `https://informationisbeautiful.net/visualizations/gender-pay-gap/`. Accessed: 2023-02-06.

[V14] Visualization number 14. `https://wien1x1.at/der-weg-zur-arbeit`. Accessed: 2023-02-06.

[V15] Visualization number 15. `https://twitter.com/Maps_US_EU/status/1472515535027978247/photo/1`. Accessed: 2023-02-06.

[V2] Visualization number 2. `http://media.economist.com/sites/default/files/imagecache/original-size/20101009_WOM993.gif`. Accessed: 2023-02-06.

[V3] Visualization number 3. `https://www.economist.com/graphic-detail/2011/12/12/the-embarrassment-of-riches`. Accessed: 2023-02-06.

[V4] Visualization number 4. https://www.reddit.com/r/dataisbeautiful/comments/iho046/how_representative_are_the_representatives_the/. Accessed: 2023-02-06.

[V5] Visualization number 5. https://www.reddit.com/r/dataisbeautiful/comments/fcgldk/how_disney_makes_its_money_oc/. Accessed: 2023-02-06.

[V6] Visualization number 6. https://viecer.univie.ac.at/blog/beitraege/2020-07-24/. Accessed: 2023-02-06.

[V7] Visualization number 7. https://informationisbeautiful.net/visualizations/food-waste/. Accessed: 2023-02-06.

[V8] Visualization number 8. https://insightfulinteraction.com/immigration200years.html. Accessed: 2023-02-06.

[V9] Visualization number 9. https://www.spiegel.de/panorama/die-welt-bleibt-zu-hause-a-5e426993-0002-0001-0000-000144545900. Accessed: 2023-02-06.

[Vis] Visual Communication - The future of content marketing. https://visualcommunication.agency/. Accessed: 2022-10-17.

[VW05] Jarke J. Van Wijk. The value of visualization. In *VIS 05. IEEE Visualization, 2005*, pages 79–86. IEEE, 2005.

[W3C] W3C Web Content Accessibility Guidelines (WCAG). https://www.w3.org/TR/WCAG/. Accessed: 2022-10-12.

[War19] Colin Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann, 2019.

[WB16] Janet Woollen and Suzanne Bakken. Engaging patients with advance directives using an information visualization approach. *Journal of Gerontological Nursing*, 42(1):16–20, 2016.

[WGK10] Matthew O. Ward, Georges Grinstein, and Daniel Keim. *Interactive Data Visualization: Foundations, Techniques, and Applications*. AK Peters/CRC Press, 2010.

[Yar19] Jillian Ruth Yarbrough. Infographics: in support of online visual learning. *Academy of Educational Leadership Journal*, 23(2):1–15, 2019.

[YWLL20] Peter ChunYu Yau, Dennis Wong, Woo Hok Luen, and Joseph Leung. Understanding consumer behavior by big data visualization in the smart space laboratory. In *Proceedings of the 2020 5th International Conference on Big Data and Computing*, pages 13–17, 2020.