CHARLES UNIVERSITY

Faculty of Arts

Department of Psychology

BACHELOR'S THESIS

David Augulis

Deceptive techniques in data visualization

Prague, 2021

Supervisor: Mgr. et Mgr. Filip Děchtěrenko, Ph.D.

Acknowledgment

I would like to express my gratitude to the supervisor Mgr. et Mgr. Filip Děchtěrenko, Ph.D., for his continues willingness to help and beneficial consultations, both well above the standard that can be expected of a thesis supervisor. I believe his guidance has taken this work to a higher level.

Poděkování

Tímto bych chtěl poděkovat vedoucímu práce Mgr. et Mgr. Filipu Děchtěrenkovi, Ph.D., za jeho neustálou ochotu pomoci a přínosné konzultace, oboje vysoce nad standardem, který může být od vedoucího očekáván. Věřím, že jeho rady posunuly tuto práci na vyšší úroveň.

Prohlášení

Tímto prohlašuji, že jsem tuto bakalářskou práci vypracoval sám a že práce v ní obsažená je mým vlastním dílem. Rovněž potvrzuji, že jsem použil pouze uvedené zdroje. Tato bakalářská práce nebyla v této podobě předložena jiné vysoké škole k udělení akademického titulu.

Declaration

I hereby declare that the present bachelor's thesis was composed by myself and that the work contained herein is my own. I also confirm that I have only used the specified resources. The present thesis has not been submitted to another university for the award of an academic degree in this form.

Prague, 24.7.2022 David Augulis

Abstrakt (česky)

Klamavé techniky zobrazování dat jsou stále více se objevujícím fenoménem. Dříve bylo tak nákladné a pracné vytvořit i prosté vizualizace, že se dalo spoléhat na autoritu jejich tvůrců. To však dnes už není pravdou a každý s přístupem k počítači a internetu je schopen tvořit bezprecedentní množství vizualizací, část z nich úmyslně či nikoliv klamavých. Vizuální gramotnost není součástí základního vzdělání, a tak nemá většina lidí, jak se proti nim bránit. V bakalářské práci shrnujeme dosavadní poznatky týkající se klamavých technik a výzkumných oblastí, které vyjasňují jak mohou být vizualizace klamavé. V praktické části jsme změřili velikost efektu 9 klamavých technik, z nichž 5 ještě nebylo testováno na souboru vysokoškolských student (N=724). Zkoumali jsme také, jak je tato klamavost modifikována poskytnutím jednorázové textové intervence, pro kterou jsme vytvořili několik úrovní, stoupajících ve své podrobnosti. Zároveň jsme taky změřili visuální gramotnost všech účastníků a pozorovali, zda nemá úroveň této schopnosti vliv na efekt klamavých technik. Signifikantní klamavý efekt se ukázal u 6 z 9 technik (aspoň mezi dvěma úrovněmi manipulace). Zároveň jsme taky vyhodnotili efekt jednorázové textové intervence jako neprůkazný a navrhujeme další kroky pro výzkum tohoto přístupu ochrany před klamavými vizualizacemi. Vyšší úroveň vizuální gramotnosti se také neukázala jako dobrá ochrana před těmito technikami, vyvozujeme tedy, že jejich rozpoznání je specifická dovednost, která musí být cíleně naučena.

Klíčová slova (česky)

Vizualizace dat, klamání, klamavé techniky, grafy, intervence, vizuální gramotnost

Abstract (in English):

Deceptive data visualization techniques are an increasingly emerging phenomenon. Previously, even simple visualizations were so costly and labor-intensive to create that one could rely on the authority of their creators. However, this is no longer true and anyone with access to a computer and the Internet is able to create an unprecedented number of visualizations, some of them deceptive, intentionally or not. Visual literacy is not part of basic education, and so most people have no way to defend themselves against them. In this bachelor thesis, we summarize existing knowledge regarding deceptive techniques and research areas that clarify how visualizations can be deceptive. In the practical part, we measured the effect size of 9 deceptive techniques, 5 of which have not yet been tested on a sample of university students (N=724). We also investigated how this deceptiveness is modified by the presence of a one-time textual intervention, for which we created several levels, increasing in their detail. At the same time, we also measured the visual literacy of all participants and observed whether the level of this ability affected the effect of deceptive techniques. A significant deceptive effect emerged for 6 of the 9 techniques (at least between the two levels of manipulation). At the same time, we also found the effect of the single text intervention to be inconclusive and suggest further steps for research on this approach of protection against deceptive visualizations. Higher levels of visual literacy also did not prove to be a good protection against these techniques, so we conclude that their detection is a specific skill that must be purposefully taught.

Klíčová slova (anglicky):

Data visualization, deception, deceptive techniques, graphs, interventions, visual literacy

Table of contents

Introduction

The current era is defined by the struggle for attention. Companies, newspapers, and politicians are spending billions to capture and maintain attention by all means available, often by trying to provoke an emotional reaction. Graphs or other visualizations are often used to support this, and for several reasons: graphs attract attention more than text, they allow large amounts of data to be summarized quickly and, finally, they have an aura of objectivity around them. It is this assumed objectivity of visualizations that lends them their ability to mislead, either intentionally or by mistake, and even major news agencies such as Reuters, CNN or FOX News are not so objective that they do not occasionally use misleading visualizations. (Pandey et al., 2015)

As Pandey et al. (2014) showed, graphs lead to higher persuasion when viewers do not have a strong negative attitude towards the message the visualization is trying to convey (if they do, charts do better). Therefore, it is not surprising that closed groups of like-minded people emerge who spread (often inaccurate) visualizations among themselves without critically evaluating them. This phenomenon is documented in detail by Lee et al. on a population of anti-maskers and shows how it can be life-threatening (Lee et al., 2021). However, this is not just an effect of shutting out opposing arguments; most people lack the ability to recognize deceptive visualizations, for example in a study in 2015, Pandey et al. shows, that 79% of participants were deceived by inverting the Y axis of a line chart and so identified the trend as descending when it was ascending (Pandey et al., 2015).

This forms another problem because charting software is widely available and tries to be more intuitive and accessible to all, which means it more often chooses the default visualization parameters itself rather than letting the user decide and as Lauer and O'Brien (2022) point out, even graduates taking data visualization courses don't feel confident in changing them, even if they recognize the visualization as being deceptive.

So, there is a great potential for a practical benefit in research of deceptive techniques and especially in how to defend against them. While the fact, that some visualization choices might mislead is nearly as old as modern visualizations (see Swoboda,H., Císař, J., (1977) or Huff, D., (1954)) , the amount of research on their deceptiveness is rather slim. This may be possibly because it is an issue that has come into the spotlight only recently, in the wake of the coronavirus pandemic, or because it is not a problem regarding scientific peer-reviewed visualizations. This is supported by the fact, that most of the research focuses on how not to be misleading or deceptive, as outlined in the theoretical part. And while important, it seems, like the current research very much neglects the existence of visualization creators with malicious intent or those creating bad visualizations by accident. In other words: it focuses on nearly exclusively on how creators can do it right, not what users can do, when they don't. A fact indirectly pointed out by McNutt et al. (2021) in their tongue in cheek work named "Visualization for Villainy", in which they summarize the existing knowledge on how to cause the most harm with visualizations, deceptive visualizations being a part of that.

Work of Pandey et al. (2015) is perhaps the most important contribution to the topic of this paper, as they were the first to measure the effect size of deceptive visualizations and so they've put in numbers, just how big a threat they pose. The techniques used were adapted from real life examples of deceptive visualizations, among them ones published by Reuters or on the official website of U.S. House of Representatives. They also come up with a twofold division of deceptive techniques: 1) message reversal and 2) message exaggeration/understatement. The logic behind it is, that each of basic graphs is trying to communicate a message which is based on the underlying data and should ideally be 1:1. If some visualization choices shift the ratio either up or down, then they must be by definition deceptive (if used in a way that shifts the ratio away from 1, if they bring it closer to 1, then we could say, that they promote understanding of the message).

To our knowledge, the work of Camba et al. (2022) is the only study focusing on how to protect viewers from deceptive visualizations. Their stance on the topic can be extracted from the name of their paper: "Identifying Deception as a Critical Component of Visualization Literacy". The authors argue, that identifying deceptive graphics is a crucial skill in the modern, visualization-filled era and should be an explicit part of visualization syllabi. They also identify 3 effective methods to do so: in-class discussion about deceptive visualization, self-learning, and peer challenge. Peer challenge is their label for an activity, where student first creates a deceptive and non-deceptive version of a chart. All the charts were then pooled and randomly presented before the class, who had to identify the deception (or its absence). After the class decided, the author explained their choice of deceptive tactic. All three of the teaching methodologies provided a significant increase in deceptiveness recognition, with success rate rising with the amount of involvement from students. The ranking of involvement in ascending order is in-class discussion, self-learning, and lastly peer challenge. The in-class discussion increased the rates of recognition from 7.62% to 35.71% (~ 4.7x increase) and for the peer challenge the rates went from 12.28% in the pre-test to staggering 92.10% in the post-test (7.5x increase).

They focused on the topic of our interest but approached it from an educational point of view; they only measured whether the user recognizes the visualization as deceptive. Whether he then correctly extracts the underlying data is a different question, one we hope to shed some light on in the empirical part of this thesis.

The goals of our work are twofold: in the first part of our thesis, we aim to introduce the reader to the topic of deceptive visualizations and the relevant findings surrounding the topic. In empirical part we expand the current research on deceptive visualization by conducting an experiment measuring the deceptiveness of 9 different techniques, 5 of them are to our knowledge yet untested. We see the main contribution of this work as shifting the focus on interventions as protection against being fooled by misleading visualizations and measuring their effect, instead of recommending good practices for visualization creators, which is the current focus. As we argue in the theoretical part we see this approach as outdated and not fitting the current broader environment. The specific research questions that we have set for ourselves are presented in the introduction to the empirical part.

1 Theoretical framework

1.1 Definitions of terms

To fully understand ourselves it is necessary to define what we mean by the terms used in our work, because almost every author comes up with his own definition. Firstly visualization: this term has two meanings for us. Firstly, as a verb it is used to mean creating a graphic (visual) representation of data, normally in order to better convey a message that comes from the underlying numbers, or to help with analyzing said data by a human. As a noun, we mean any result of this activity; for our work (especially the empirical part) you can think of the most common graphs we encounter on a daily basis: a pie chart, a line graph or a bar chart. Although these types of graphs are far from exhausting the content of the term data visualization, for example, 3D modeling of real objects, displaying forces on an object, or simulating trajectories are all data visualizations too, but they are not useful for the purposes of our work, so we can exclude them. So by visualizations we will mean purely graphs and charts.

A deceptive technique is any choice of data visualization method that leads the user to incorrect conclusions about the form of the underlying data. There is a plethora of such techniques, and we will discuss their analysis and classification in the theoretical section. Each particular technique can also be called a manipulation, for example cropping the Y-axis to make the difference in the data appear larger is one type of manipulation. At the same time, we will usually not be interested in the intention behind the creation of deceptive techniques. Let us use the Merriam-Webster Dictionary's definition of deception: "the act of causing someone to accept as true or valid what is false or invalid."(Deception. In Merriam-Webster.com dictionary, 2022) This definition does not presuppose intent or knowledge that we are conveying an untruth, just as many authors of deceptive visualizations may have no idea that their creation is misleading. Although for some deceptive visualizations (especially those with political agendas) it is not difficult to reverse engineer their intent.

Closely related to deceptive techniques are the levels of manipulation. In fact, we can truncate the Y-axis by almost any number of units, we can truncate it by 10% or we can truncate it by 100% (i.e., make the data fill the whole graph), so by levels (of deceptive techniques) we understand different forms of same technique. In our example truncating the Y axis by 10% would be one level and by 100% a different one, but they all fall under the same manipulation.

We will also be talking about visual literacy. In the empirical part of the thesis, we will use the term visual literacy (sometime VL or VL score) to mean the numerical score that each proband receives as a result of the test by Boy et al. This test measures how correctly a person can read values from tables and line graphs in a predetermined amount of time. In the theoretical part, it will mean a similar concept, but we will not be quantifying it. It represents a broad ability to understand, read from, and create visualizations. John Debes, the founder of Visual literacy association does not reduce the concept of VL strictly to data visualizations but applies it to all visual stimuli, be it visible actions, objects, or symbols (Debes J., 1969). This definition is too broad for our purposes, although we do not claim the right to a final definition of this term, which has been considered controversial and unclear since its inception (Avgerinou & Ericson, 2002), and still waits for a broadly accepted definition, as each

1.2 History of datavis and why is it a problem

The first notable use of graphs as we know them today was in a revolutionary book by Scottish engineer William Playfair. A reader of his book, named The Commercial and Political Atlas, was probably surprised, that an Atlas contained no maps, but in a way it did. Playfair's breakthrough idea was: what if instead of using latitude and longitude as we do for maps, we marked different variables on the vertical and horizontal axes? After all, it is all a measure of some quantity. For example, we can switch longitude in degrees for time in years and latitude for the amount of goods as shown in ***Figure 1.*** Each vertical line

would then mark a 10 years difference and each horizontal line a 10 * 10,000 pounds jump. We call these smallest distances "tics" and their labels on the axis "tic marks".

With such a system Payfair then marked the amount of imports and exports for each year and connected substituted the dots for 2 lines, the area between exactly quantifying the sum of differences for each year. Nowadays we recognize this as a simple line chart, but in year 1786 it was anything but. The instructions on how to read it and what it shows spanned multiple pages.

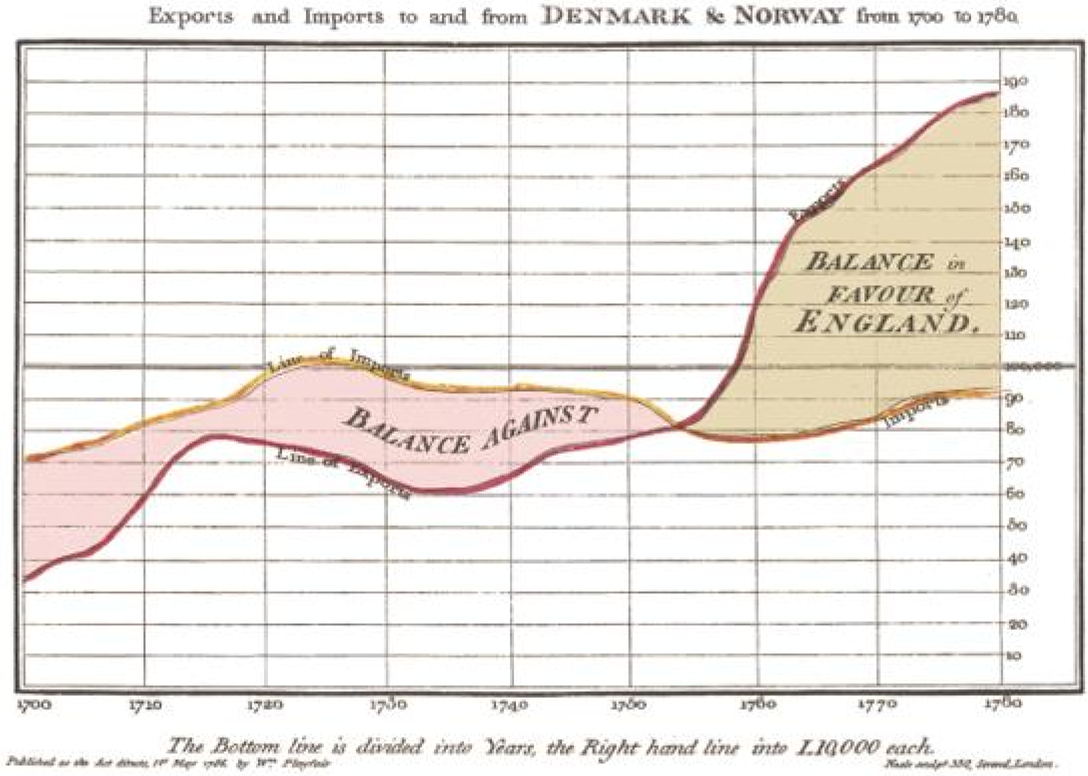


Figure 1- An early line chart by Playfair

This wasn't by any means the first visualization of data, but it was the first meaningful visualization of quantitative data in a chart (Few, S., 2012). So Playfair was the one started visualizing the data and not only using line charts, but already using a pie chart to show the ratios of groups in the whole.

This information may seem like piece of trivia, not important for the topic at hand, but when he started the trend of visualizations most visualization creators who came after him stood on his shoulders so to speak and used his creations as the basis for their own (Few, S., 2012). The problem with that is, that at the time of publication, it would be another hundred years before the first laboratory of psychology is founded by Wundt, and a half of another century before anyone applies the scientific method to research, whether the chosen visualization techniques are conveying the underlying data without any systematical biases or other interferences. And this issue persists till today, simply said: we use many data visualization techniques and principles simply because we have always used them, not because we found them to be the best (Kosara, 2016). As Kosara points out in the same publication, fittingly named "An empire built on sand", if we were to design datavis from scratch today using what we know about perception, cognition and human-computer interaction, we would most likely end up with a different system. Instead of building on what we know, we first had the system in place and then started testing out some of its components. Undoubtedly some methods we now use would remain, because they work as we want them to, but the catch is, that we do not know which they are until we test all of them. An example to illustrate this point: imagine seeing a graph shown in **Figure 2.** in a scientific or academic publication.

Figure 2- An ISOTYPE bar chart

A person not familiar with the latest findings would probably think, that the author is either joking or trying to discredit himself. At the very least one would think that the choice of encoding is very non-standard, and it is. This type of encoding is classified under an umbrella term called "visual embellishment" (or chart junk for the uninitiated (Tufte, E. R. (1983))). This exact case is an "ISOTYPE" embellishment, as it is using a small pictographical elements instead of standard bar and is considered a suboptimal choice by many. It is almost impossible to find in any serious publication nowadays, but without any evidence suggesting it is in fact a bad choice for encoding length (quantifying effect by area of these pictograms is a different matter, but that is not because of any inheritant issue with this manner of visualization, but because encoding effect as area simply does not mesh well with our visual system (Franconeri et al., 2021), as we will discuss later). But recent findings show that this discrimination is not supported by evidence. On the contrary, using cats and parrots instead of the usual bars increased recall of information presented as well as user engagement (Bateman et al., 2010; Haroz et al., 2015). So in Tufte's book The visual display of quantitative information (one of the most influential books on visualizations, with upwards of 3200 citations) we can read not to use so called "chart junk", even though it has little to none undesirable effects (when used as instructed), but we can also read about deceptive techniques and their misleading effect on user, and this assumption for a change has been found to be correct (Pandey et al., 2015).

Point of this chapter is to illustrate, how even though we would like to think of datavis as an exact discipline standing on a strong body of research, it unfortunately isn't so (just yet). Even though deceptive techniques are only a small subset of the field of data visualization, the entire field itself is so far largely based on historical precedent and convention rather than verified facts. That is why in the empirical part in addition to confirmed deceptive manipulations we chose to investigate some techniques, which are not currently believed to be deceptive, but to our knowledge no one has checked if that's really the case. Even if we find them to not to be misleading, we at least create a small support point from which we can continue to reinvent this empire built on sand.

1.3 The benefits of visualization

To fully appreciate the effect of deceptive visualizations, we deem it useful to first briefly summarize the significance of visualizations. After all, the fact that some visualization techniques are deceptive would not be a problem if the visualizations themselves had no effect on society or the individual. The issue is that we do not know that they are deceptive and so they can hijack the powers that visualizations possess.

1.3.1 Providing insight

The human brain is by no accident called the most complex system in the universe. And this system, so complex that we are not even close to fully understanding it, let alone reproducing it, is almost half made up of the visual subsystem, the largest single system of the entire brain (Van Essen et al., 1992). And although made for orienting us in the wild and simplifying the infinitely complex and ever changing environment we live in (What Is the Bandwidth of Perceptual Experience?, 2016), the visualization language of shapes and colors taps into its computational power rather well by processing the data parallelly (the stimuli don't compete with each other, are evaluated independently of each other) instead of serially (the way we process numbers out of a table for example, we need to focus our attention at one number at a time). We can "digest" vast amount of data this way and immediately identify the trend and outliers with nearly no conscious mental effort (Franconeri et al., 2021). While visualizations and numbers are both artificial means of communication, the former offers us greater bandwidth for transferring information because it utilizes evolution's most complex computational system.

1.3.2 Persuasion

The persuasive aspect of visualizations is particularly interesting, even more so in the context of our work, because of how deceptive visualizations might hijack it. One might think, that the same data should not have different power to change one's opinion based solely on the form of its display, yet it does. Elting et al. (1999) reports, that the choice on how to represent data concerning a clinical trial had significant effect on whether the person in charge decided to continue funding the trial or decided to shut it down. Reporting individual improvement of each patient was more persuasive to keep the trial running then providing the same data in form of a table or visualizing the progress, but just as a summary of the whole set of patients.

Tal & Wansink (2016) report a similar finding. Their study showed that participants deem a made up drug as more effective, if its efficacy is reported in the form of bar chart instead of plain text. They call this effect being blinded with science. A replication study however found that the effect might have been overstated and the bar graph does not persuade, but rather allows readers to better evaluate the effects of the drug (Dragicevic & Jansen, 2018).

This is perhaps the greatest danger of deceptive visualizations. The aforementioned studies in tandem with study from Pandey et al. (2014), which shows, that people are more often persuaded by visualizations, that shows them a message, with which they do not strongly disagree give us an insight into why we so often see shared blatantly misleading visualizations (most often with social or political topics). This phenomenon has not escaped the attention of the general public, and accounts are emerging that collect similar misleading visualizations (see www.twitter.com/GraphCrimes or www.viz.wtf)., mostly for entertainment purposes

1.3.3 Memorability

While memorability of visualizations is of lesser importance then the two previous effects, it still offers some interesting insights. For example, it is documented, that visualizations featuring bold colors or humans are generally more memorable than those which don’t (Borkin et al., 2013). The same study concludes that unique visualizations are more memorable, than common plain visualizations and the authors advocate for the use of aforementioned “chart junk” as a way to increase recall.

Marriott et al., (2012) focuses their research on layout features such as symmetry, collinearity, or orthogonality. Users were shown simple objects made of lines and circles as shown in **Figure 3** and were later asked to draw them from memory. Their findings show that those three aspects of graphs make them the most memorable out of all the features studied, some features that were not proven to be significant in modifying memorability are node alignment or the use of parallel lines.

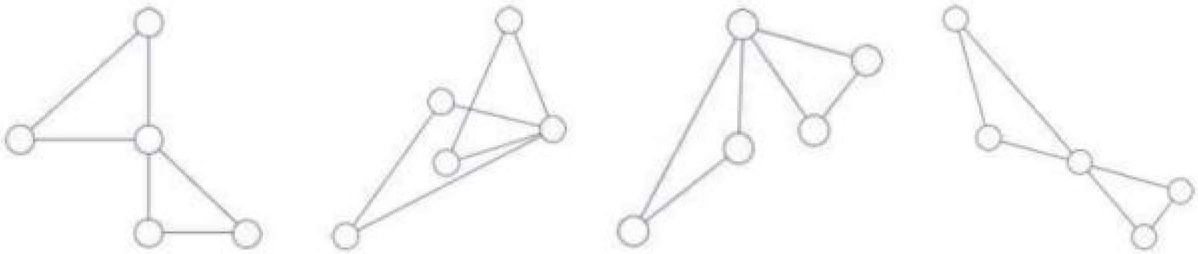


Figure 3- Objects used in the study, each one created with set features such as symmetry or parallel lines

1.4 How can visualizations mislead

Now that we summarized what features of visualizations the deceptive techniques might hijack, we can finally list some of the most notorious ones. It is important to remember, that the term deceptive technique, as we used and defined it earlier, captures only a small portion of the underlying substrate of deceptive visualizations, which is itself again a small part of how graphs can be harmful or misused. For example, McNutt et al. (2021) uses two dimensions to categorize all the ways visualizations can be used to cause harm, some of which we probably wouldn't have thought of.

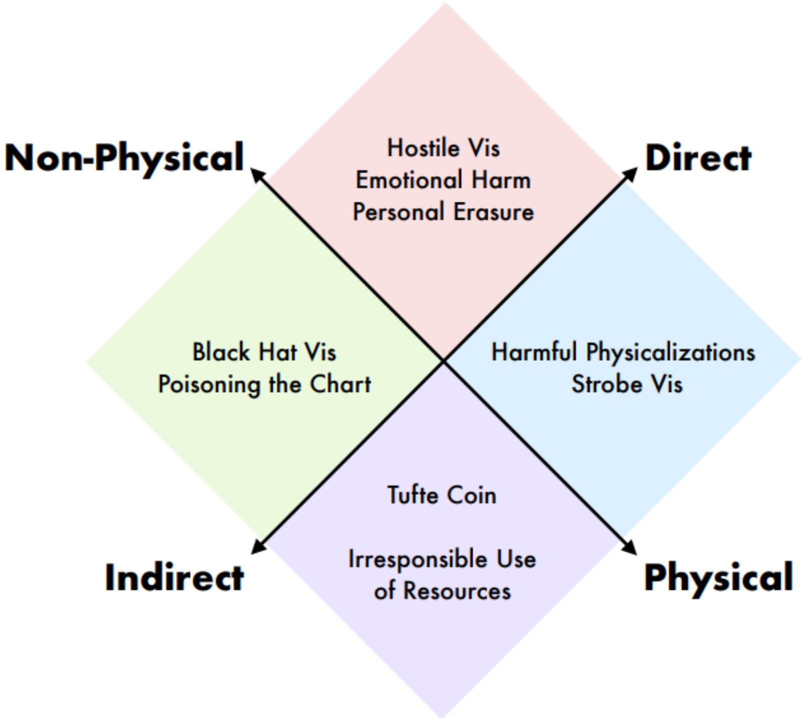


Figure 4- the 2-dimensional model of harmful visualizations from McNutt et al. (2021)

As we can see in **Figure 4**, we can divide harmful visualizations into 4 categories depending on their position on two spectrums*: physical - non-physical* and *direct – indirect.* We must say, that visualizations that cause physical harm are more of a thought experiment rather than any real menace to the society at the present moment. They involve inventions such as creaing a "Tufte coin" cryptocurrency, named after the already mentioned Edward Tufte a pioneer in the field of datavis, which would connect visualization to blockchain and by using it, the users would unknowingly (or knowingly) waste valueble resources (electrical energy and the means of creating it), which would be more needed elsewhere. The second is "Strobe Vis", which advocates the use of quickly flashing lights to induce seizures in viewers and reduce accessibility.

The most relevant harmful visualization for the goals of our work is in the indirect – non-physical quadrant and authors named it *Black hat* vis, in reference to term used in cyber security, where a person with malicious intent is called a black hat. They classify such offences against the correctness of visualizations as usually involving a "man in the middle" attacks, meaning that data coming in is correct, as well as the intention of the viewer, but they are separated by an actor with malicious intent, for example a data visualizer, who wants the data to be inaccessible for the viewer or wants him to draw the wrong conclusion. These attacks can be performed by the implementation of tactics demonstrated by Pandey et al. (2014) such as inverting the Y axis to make the trend seem as going in other direction or choosing a poor graph type for the query at hand.

But classifying all uses of deceptive techniques as black hat would be a mistake, because it expects a malicious intent of the creator, which is not always the case, so they produce another spectrum, ranging from “stupid hat visualizations” to “black hat visualizations.” While the result may be the same, a design falling in the former category was not made intentionally deceptive, but came out as such because of poor design choices, usually caused by lack of author’s knowledge about datavis. We believe that the “stupid hat” visualizations make up a non-trivial part of the body of deceptive visualizations shared on the internet. In the next subchapter we will summarize the most common deceptive tactics, featured in both stupid hat and black hat visualizations.

1.4.1 Common deceptive techniques

Perhaps the most prominent deception technique is manipulating the Y axis in some manner, whether it is changing its aspect ratio (making the trend or effect seem larger or smaller) or showing incorrect interval for the task at hand. They can be used for line chart, bar chart, scatterplot and many others that rely on the classical perpendicular x-y axis set up and code data by either position or length. These tactics have a remarkably high potential to be deceptive, because we are very skilled at comparing lengths and positions (Franconeri et al., 2021) and so this precision and our subsequent trust in the results lend this technique its deceptive potential.

Older books on this topic (see Swoboda,H., Císař, J., (1977) or Huff, D., (1954)), often claim, that truncating the Y axis a visualization sin and should not be done, unless absolutely necessary (they mention a lack of space on page for the visualization as one of the possible reasons for using it). But current publications contradict this opinion. For example, Correll et al. (2020) point out, that depending on the task, **not** truncating the Y axis to fit the dataset can be the deceptive choice. In **Figure 5** we see the same data visualized line chart with non-truncating Y axis (left) and truncated Y axis to fit the data (right), showing the development of the average temperature in Fahrenheit for each year from 1850. As we can see, showing 0 in this type of graph for this type of task is the deceptive visualization, as the slight differences are almost non-visible. On the other hand, focusing only on the narrower ranges allows us to fully appreciate the growing trend.

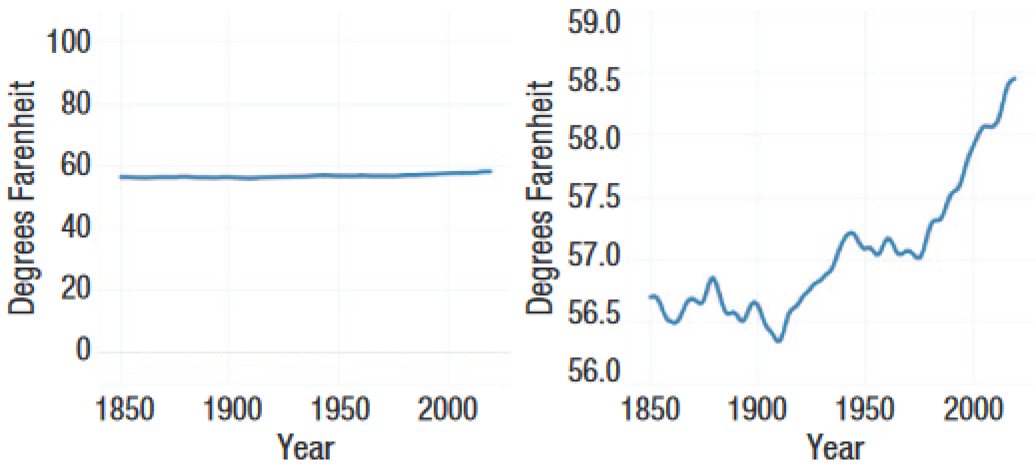


Figure 5- A case for truncating the Y axis, original source Franconeri et al. 2021, p.117

Pandey in the same study finds an interesting effect, the users are still influenced by truncating the Y axis (they report the effect as being larger than it is in the underlying data) even when they are clearly made aware the Y axis has been modified. That creates another research topic for the science of datavis: what is the effect deceptive techniques on users, when they are made aware of their presence in the visualizations they are seeing. The short answer: we do not know, to our knowledge no study has yet explored this problem.

A terrific way to thoroughly mislead an inattentive reader is a tactic called Axis Inversion. Simply put the visualizer goes against the convention and does not make the axis in question start at 0, but the other way around as is shown on **Figure 6**, a real-life example of this technique published by the news agency REUTERS.

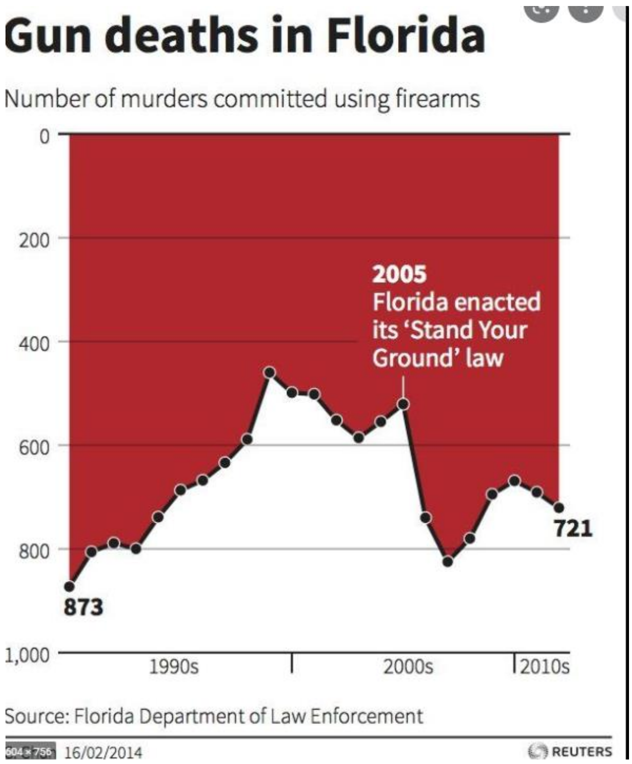


Figure 6- A chart featuring the inverted Y axis (used in Pandey et al., 2015, p. 1470 and others)

Pandey et al. (2015) measured its effectiveness and found, that participant who were shown a similar graph answered incorectly in 78.95% cases, when asked about the direction the trend is headed. For comparison that number was just 2.5% for the non-deceptive version. Although it is usually the Y axis that is inverted, it is possible to do it for the X axis as well.

The last deceptive technique we would like to describe is concerning a bubble chart and plays on the mathematics of different rate of increase for area of circle and the diameter of circle. It is encoding the effect as diameter, instead of the usual area and is presented on **Figure 7**.

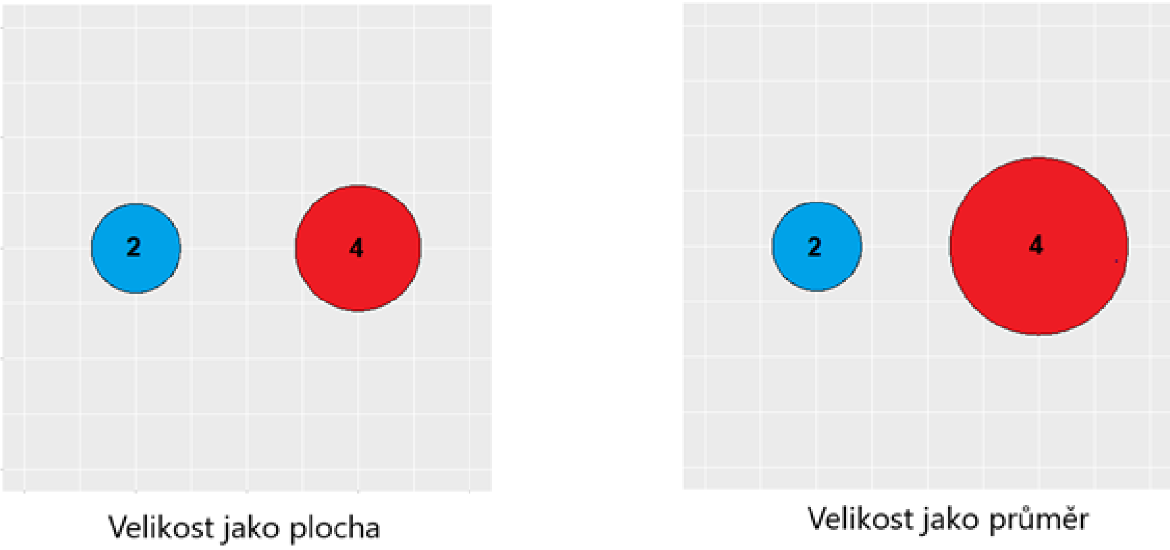


Figure 7 effect as area (left) and as diameter (right)

The graph is a part of textual intervention we created, to warn participants in our study about the deceptive effects of this visualization. Because of the math involved, if we increase a diameter two times, the area of the circle increases by the power of 2, the effect gets more distinctive for larger numbers (if we increase diameter 4 times, the area increases 16 times). If the creator a graph does not include the underlying numbers, it is hard not to get deceived, because it is impossible to know, whether the author used area or diameter (or even radius) to encode the effect. Pandey reports this manipulation as having significant effect, choosing to deploy this tactic increased the mean answer on 5-point Likert scale by 1.00 (from M=1.71, 95% CI [1.45, 1.98] for the control to M=2.71, 95% CI [2.27, 3.15]).

Those are some of the well known and researched deceptive techniques, but just to name them all is beyond the means of any bachelor thesis.

1.4.2 How to recognize deceptive modifications of graphs

Tufte in 1983 produces a way to quantify the deceptiveness of a visualization by a simple formula:

$$\text{Tufte's Lie Factor} = \frac{\text{size of effect shown in graph}}{\text{size of effect in data}}$$

Where what we would call a non-deceptive visualization has a lie factor of 1, meaning the graphical representation of the effect, as we can physically measure from the final graph must be equal to the effect size in the numerical data. Deviating from 1 in either direction would be a violation of Tufte's law and therefore mean, that said visualization is misleading. In his publication he states that any deviation greater than 0.05 in the lie factor indicates a significant distortion. (Tufte, 1983).

While there is a discourse about this metric; some think it was made to poke fun at the rigorous movement of the time of trying to quantify everything (Few, S., 2012), and some used to swear by it and compute it for vast number of graphical representations (Jarvenpaa, S. L., 1988).

We see this as a reductionist metric that does not cover all relevant aspects of misleading visualizations (e.g., it completely fails to recognize overwhelming the user with a lot of unnecessary data so that essential relationships remain hidden), but it can be interesting to see, if the deceptiveness of a manipulation increases proportionally to the increase of Tufte's lie factor. Also, it can serve when reading a chart, the reader with knowledge of this law can always ask a question: "is the effect not presenting itself to me as larger/smaller than the data indicates?", for example by trimming the Y-axis or changing the ratio of the axes.

Cairo (2014) puts forth his own checklist on the 3 ways visualizations can be deceptive, they are:

1. Hiding relevant data to highlight what benefits us
2. Displaying too much data to obscure reality
3. Using graphic forms in inappropriate ways (distorting the data)

In the empirical part we will focus on the third category, but we can appreciate the deceptive potential of the other two as well. For example, a very skillful way of using too much data to make any relevant information unattainable is a flowchart, made by the republican party as an argument against the health plan law of the opposing political party, as shown in **Figure 8**. While not showing any falsehoods per se, the amount of information is disproportionate to the task at hand, that is to decide whether or not passing this law will have a beneficial effect on the individual and the nation. It instead gives the impression that the new system is incomprehensibly complicated and that if the delusional democrats pass it, an honest American won't even know, how to find his healthcare practitioner.

Camba et al. (2022) comes with their own checklist, but this time for their students to go through when faced with a new visualization. The four deceptive criteria as follows:

1. incorrect type of graphic
2. incorrect range or scale
3. incorrect use of the semantic variables
4. incorrect labeling of the displayed information

They should be answered in a fail/pass manner, with only a visualization that does not fall into any of these categories can be accepted as non-deceptive and further used.

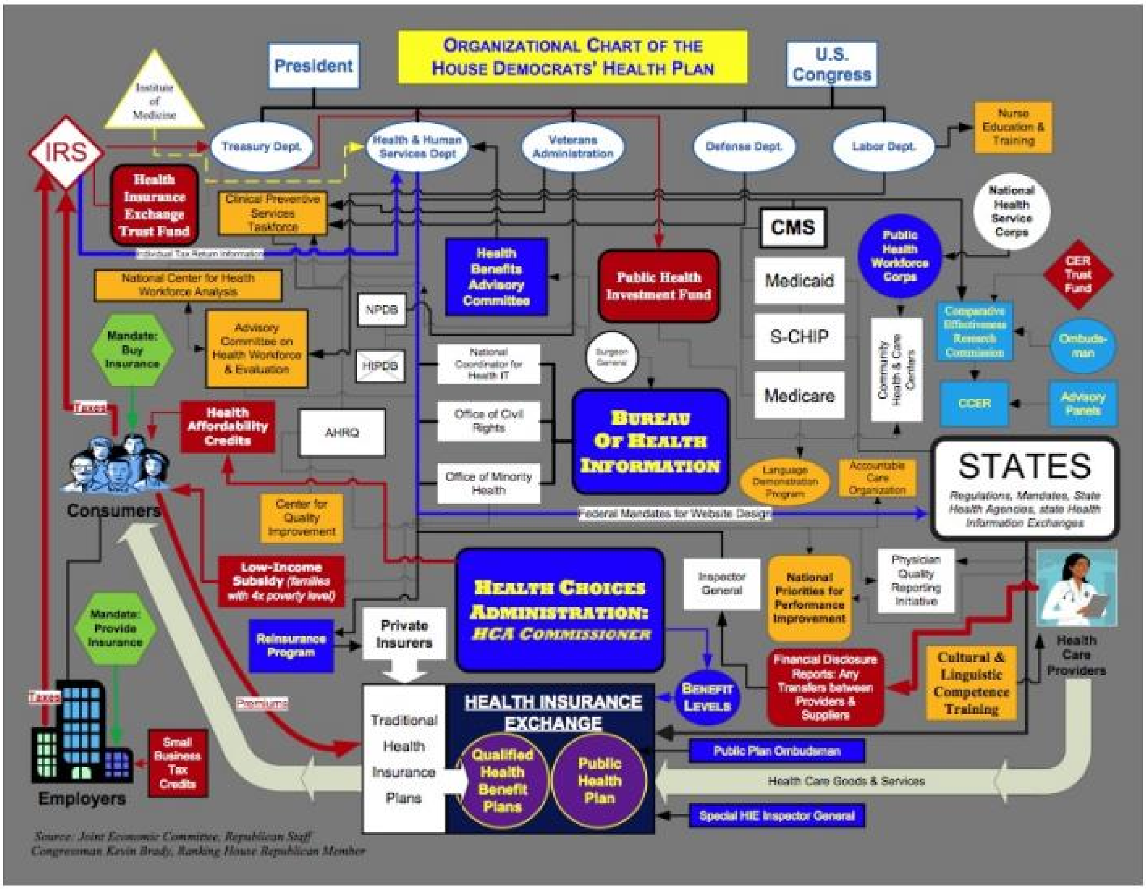


Figure 8- Organizational chart of the house democrats' health plan. Source: http://voices.washingtonpost.com/ezra-klein/assets_c/ 2009/07/jecchart.html, 2009.

1.5 Visual literacy

The Increasing visual literacy may be one solution to the problem presented. It is not controversial to say that most study programs do not prepare their graduates for the realities of modern times. Cairo (2019) proposes adding several subject areas to the syllabi of visual journalists and other professions dealing with information design. Among the subjects are for example cognitive biases, introduction to statistics or findings of visual and cognitive sciences (as relates to creating visualizations). Her argument is, that mistakes and deliberate deceptions are inevitable in the age of freedom of expression (and unprecedented options of creating and distributing these expressions to other people). And that the way to go is to create a generation of better prepared, evidence-driven visual communicators, whose main objective will be to present information without any biases or deceptions. This is one approach, he calls it "Fighting noise with knowledge."

Second direction is one Camba et al. (2022) support in their work. It is increasing the visual competence of their students to teach them to recognize deceptive visualizations. Their educational methodology has proven successful, their lowest level of educational program has yielded increase from 7.62% to 35.71% of recognized deceptive visualizations and their highest increased the percentage of detected misleading visualizations from 12.28% at pretest to 92.1% at post-test. They call the latter intervention program a "peer challenge" and it was an in-class activity, in which each student was given their own dataset and had to display it non-deceptively and then try to trick their classmates using some deceptive technique. The effectiveness of this intervention is really remarkable and in our opinion has a lot of potential to be included in a future syllabi where it is desirable to reduce the effect of misleading visualizations on students, which one could argue is nowadays everywhere.

2 Empirical part

2.1 Introduction

The empirical part of this thesis builds on and expands the work of Pandey et al. (2015) who measured the effect size of deceptive techniques. We used four of the same techniques and five others that may have a deceptive effect but have not yet been tested. Camba et al. (2022) has shown, that it is possible to teach students to recognize deceptive visualizations. The ability to do so increases with the level of involvement of the participant in the intervention; the more work the participant must do, the more likely he is later to recognize that a deception has taken place. Their lowest level of intervention was in-class discussion. We are interested in how an intervention with even lower level of involvement will help the viewers. Also, we won't measure the number of deceptive techniques recognized, but how it reduces their effect, because one might recognize that a visualization is deceptive and yet still walk away with the wrong conclusion.

We made up to three levels of adjustment for the manipulations (instead of just comparing the non-deceptive graph against one level of deceptive technique) and in addition to demographic data, we also measured participants' visual literacy using a test designed by Boy et al. (2014).

We have chosen the following research questions:

RQ1: How deceptive are the chosen deceptive techniques? Particularly how is the effect of deceptive techniques related to the level of manipulation.

RQ2: Does providing more detailed warning about deceptive visualizations lower their effect on viewer?

RQ3: How does the level of visual literacy relate to the extent to which a viewer is deceived by deceptive tactics?

RQ4: Is there any relationship between individual differences and the effectiveness of deceptive techniques?

2.2. Methods

2.1.1 Participants

Data from 724 participants aged 18 to 53 years (M = 22.8, SD = 4.9) was used for the final analysis. Out of the participants who completed both parts (N=772), we excluded those who answered the attention check question incorrectly (N=46) or had color blindness

(N=2). This is because healthy or corrected vision (without color blindness) is a prerequisite for obtaining correct results from the visual literacy test used. 15.6% (N=113) of the sample were male, 84% (N=608) were female, with 3 participants choosing the option “other” when selecting their gender. 18 participants or 2.5% selected that their mother tongue is not Czech, but that wasn’t a reason for disqualification from the dataset, as the questions were not linguistically complicated and since they were presumably studying a graduate or undergraduate program in Czech, we would expect that their language ability is more than sufficient for our research. 78% (N=568) reported their current highest education attained as high school and the second largest group was graduates with bachelor’s degree (18.6%, N=135). Curiously 1 person reported as having only finishing elementary school, which we must think was a mistake when selecting, but this too was not a reason for disqualification.

These were students at Charles University from different faculties who enrolled in an elective course in which they are awarded two credits for participation in four experiments. Participants signed up for the experiment on their own and completed it on their own devices , with instructions given to them not to use mobile devices to ensure correct display of the stimuli.

2.1.2 Choosing and Creating Visualizations

Building on Pandey's work, we focused only on simple visualizations that most often appear in the media, schools, or annual reports. Thus, we excluded complex visualizations that, although having the potential to mislead (perhaps even more so, but that is due to their complexity and inaccessibility to the untrained, which is not the focus of this thesis) and chose for our experiment following graph types: line chart, pie chart, bubble chart and bar graph. For each deceptive technique we came up with up to 3 levels of deception, where Level 1 is always the non-deceptive graph without any modification, that we used as a baseline. For some manipulations we assume that the higher the level of modification, the greater the deceptiveness (for example changing the aspect ratio to 1:2 should be more deceptive than changing it to 1:1.25) and for some it is just to differentiate between the levels used (for example pie chart vs. bar chart). The complete list of techniques and levels for each one is in **Figure 9**. Manipulations that we expected to be more deceptive with higher levels are written in yellow.

Manipulation \ Level	1	2	3	4
Truncating Y axis	None	By 25%	By 50%	By 100%
Aspect ratio	1:1	1:1.25	1:1.5	1:2
Quantity as area/ diameter	Diameter	Area		
Inverted Y axis	No	Yes		
Pie chart- rotation	By 0°	By 90°	By 180°	
Pie chart- sorting	Baseline	By area		
Pie chart vs. bar charts	2D pie chart	3d pie chart	Bar graph	Stacked bar graph
Cumulative graph	No	Yes		
Logarithmic graph	Both axes linear	X axis logarithmic	Y axis logarithmic	

Figure 9- The table of used deceptive techniques

Each non-empty cell represents one graph created, so twenty-six in total. We used a new baseline (level 1) chart for each manipulation, even if it was for the same graph type. For example, *Truncating Y ax*is and changing *Aspect ratio* are both deceptive techniques used for a line chart, so the level 1 could be the same for both, but to avoid interference we used separate graphs with different data and context.

Graphs were created in R with package "GGPLOT2" and had the same color scheme and size (except for the *Aspect ratio* manipulation). A sample of the graphs used in the study is shown in **Figure 10** and the full set of all stimuli used is available at osf.io/k59zq/.

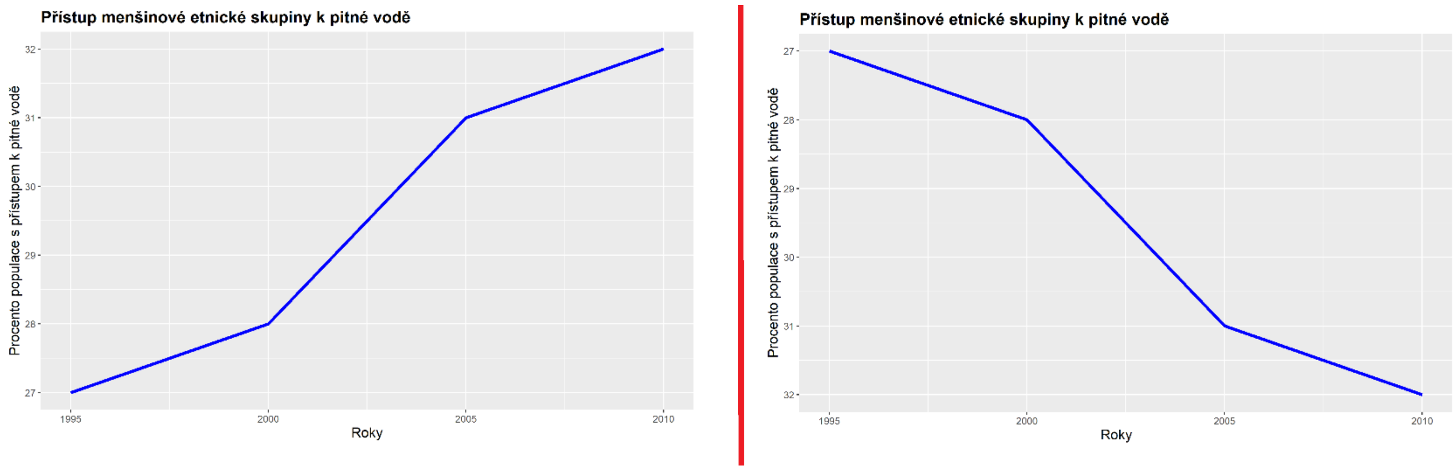


Figure 10- The graphs used for the "Inverted Y axes" manipulation

2.1.3 Interventions

To answer our second research question (*Does providing more detailed warning about deceptive visualizations lower their effect on viewer?*) we crafted 5 messages about deceptive visualizations with each with increasing level of information given. Users were randomly assigned intervention group 1 to 5, which decided what message will be shown to them. Group 1 was control group, which received no information about deceptive techniques. Second group was told that there are visualization choices, which might inhibit, exaggerate, or reverse the message of the data. The third group was told the same but was also told how that might work for 3D pie chart (how due to perspective the more distant section might seem smaller than the closer section). The intervention group 4 was also shown a 3D pie chart to better illustrate the point and the fifth group was shown both levels of a deceptive technique used in the study: displaying the effect as either diameter or area. All the used intervention messages are in **Attachment 1**.

2.2 Design

The design of the experiment was divided into 2 parts. The first part was run on the open-source survey network Formr, which is built on R and offers wide customizability of both survey pages and the whole run. The second part was measuring the visual literacy with the "Line graph 1" subtest developed by Boy et al. and was hosted on a private server. We decided to use just one of the two tests developed, because using both would have been too time consuming for the participants (one subtest took ~15 minutes to finish) and would likely have resulted in increased test abandonment or random answering and we chose "*Line graph 1*" over "*Line graph 2*" test, because they have nearly identical psychometric qualities, but the former offers slightly better informational value for users with just below average scores, which we expected might be the case in our population, which is mostly untrained in visualization.

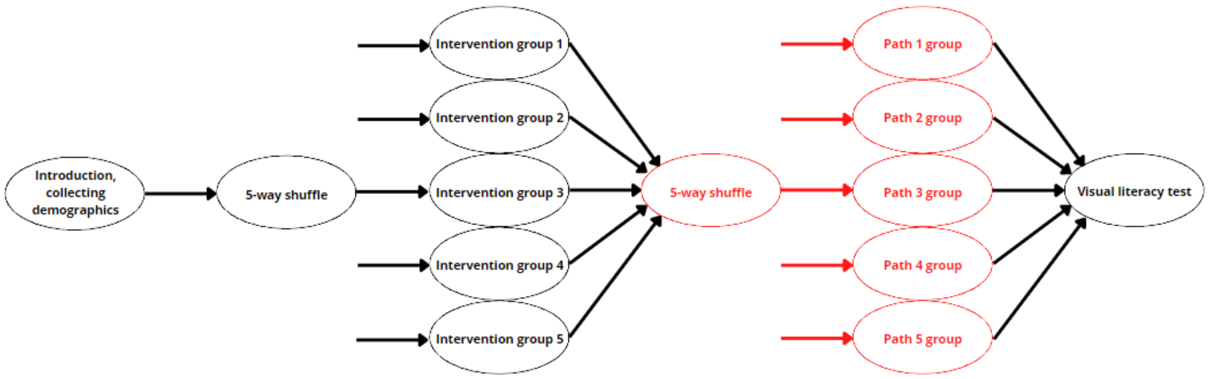


Figure 11- Diagram of the experimental design. Unfortunately, due to unforeseen inner logic of Formr framework the second shuffle was not registered by the program and so the groups remained the same. The problematic parts are marked red.

We present the study's design in **Figure 11**. The first information the probands received was in the email with the invitation to join the study. It simply stated the amount of time they need to set aside to complete it and that they will need a computer or some other electronic device with large screen, this was mainly to ensure, that they see the graphs as intended and will not need to scroll or zoom in, that would introduce intervening variables into the data, which we could not control. The invitational email also did not mention anything about deceptive techniques as not to influence those participants who will be later assigned into the control group and should not be informed about them in any way. It merely stated the topic of the research as data visualization. The first part was a questionnaire where we collected the participant's name (used for pairing answers from the first part with the visual literacy test results and then deleted, data collected was analyzed without the participants' names), gender, age, whether Czech is their first language, what is their highest acquired education, whether they think mostly with words or with pictures and if they have any kind of visual impairment. We used the question "Do you think mostly with words or pictures?" to see, whether the results corelates with the result of the visual literacy test. Pandey at al. (2015) used the same question as a proxy for the visual test, so we wanted to see, whether this approach had any merit to it.

If they chose the answer "yes" on the visual impairment question a text prompt with a line for their answer appeared, asking them to specify, as we only could use data from participants with healthy or corrected sight. The test ended for those that marked "colorblindness".

After this survey, the program randomly assigned users in a group 1 to 5 and showed them corresponding intervention message. After that should have followed another 5-way shuffle, which would separate the users once again into 5 different groups and would decide in what order they would see the 26 graphs. The paths we chose can be found in the OSF directory. They were created to reduce any effect of seeing the visualization in a particular order. Unfortunately, this did not happen, as Formr does not rewrite the first shuffling by a second shuffle in the same run, it will still only consider the first one. So, all participants from Intervention group 1 saw the graphs in the same order, same for group 2 and so on. This was a methodological mistake which might have influenced the results, we discuss it in detail in the *Discussion* portion of the empirical part**.**

There were two types of questions in the study, inspired by Pandey et al. (2015). For *Inverted Y axis* manipulation, we asked, "*What can you say about XY?*" specifically "*What can you say about the access to drinking water?*". The options were "Decreased", "Not sure", and "Increased", so we were measuring whether the participant correctly recognizes the direction of the trend. For the other manipulations we asked "**How?**" type of questions. For example: "*How do you assess the population growth trend over this time period??*" for the *Logarithmic graph* manipulation.

As the 20th question in each run, we showed the participants an attention check. In it, instead of a "How?" question, we asked to mark the "1 - Minimal" answer. We chose a bubble chart for this because the user might be able to guess what we wanted to ask from context of previous questions, when presented with a line chart, but for a bubble chart it is impossible to answer correctly without reading the question. The attention check is shown in **Figure 12**.

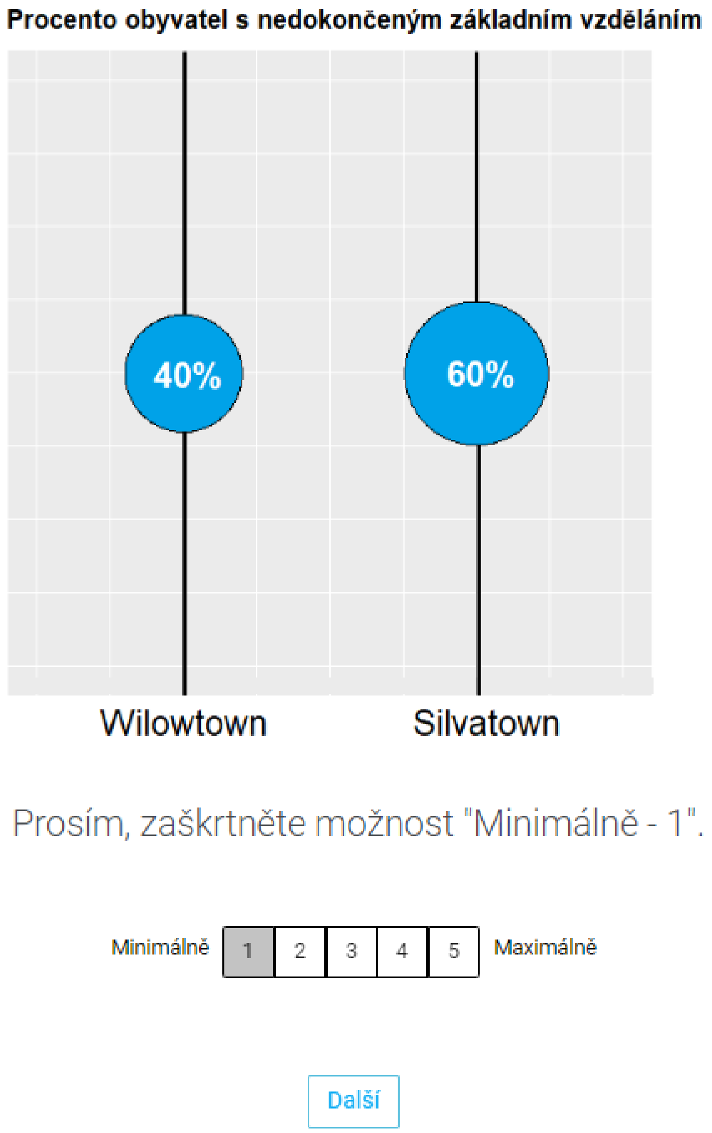


Figure 12- Attention check as used in the study. Translation: Please, select the option "Minimum - 1". In bubble graph the participants compared two values against each other and so needed to know which one they should serve as a baseline for the comparison, so reading the question is a necessity for this type of graph. That is why we chose it for the attention check.

After the participant saw all questions and deceptive visualizations, they were redirected to external website we have set up, where we hosted the visual literacy test. We did not change the test in any way, just translated it to Czech. Also, although the probands were redirecting immediately after filling out the first part, they were not forced to start right away. Users had the option to fill the second part whenever they wanted (in the 7-day time window we have set up). After they finished the visual literacy test, they were once again asked for their name and upon filling it out their answer was saved to a document on the same private server. The server was password protected and no-one had access to it except for the experimenters. When the phase of data collecting ended we closed both parts of the study, used the names to pair the results from both parts and then deleted them.

Participation was entirely voluntary and participants could withdraw from the study. Participation in the study carried no ethical risks. Participants' data were password-locked at all times and only examiners had access to them. After the data from both sections were paired, the data were anonymized.

2.3 Analysis

RQ1:

As our experiment is mixed design with multiple comparison groups and a potential within-single-subject covariate, we will analyze the data with analysis of variance with the following arguments:

dependent variable: participants' answers to each question

within-Ss predictor: level of manipulation

between-Ss predictor: intervention group of participants

within covariate: score in visual literacy test

Firstly, we will analyze the data without the visual literacy score, then add it as a covariate and compare the results. We will also deploy Mauchly's test for Sphericity on data from techniques with more than 2 levels of manipulation to check, whether the sphericity assumption isn't violated (this will be reported as NA on techniques with just 2 levels, as sphericity holds for 2 levels). If the assumption of sphericity is not met (the Mauchly's test comes back as statistically significant), we will perform sphericity corrections using the Greenhouse-Geisser and Huynd-Feldt corrections. We will also calculate adjusted p value using the Holm–Bonferroni method to reduce false positives in our design with multiple comparisons. Effect sizes will be reported in generalized eta squared, as for our purposes the statistical significance is not as interesting as real-life application, and in our rather large dataset with multiple comparison can be misleading. We will also do a post-hoc analysis and visualize the findings, splitting groups by intervention, as it is a factor of interest and averaging the score for all interventions might hide interesting information. Lastly, we will calculate estimated marginal means for each level.

RQ2:

To answer this question, we will use results from the analysis for RQ1, inspecting the effect of intervention and its interaction with levels of manipulation. Pairwise comparison of EMM for each intervention group will be reported.

RQ3:

We will calculate the difference between level 1 and the other levels for each manipulation that proves as deceptive and be left with a difference in answers for each level for each participant. This number tells us, how much the participant changed his answer for depending on level of manipulation. We will refer to this number as "Deceptiveness Delta (DD)". We will examine correlations between DD and VL score, we hypothesize, that they will be inversely proportional for each manipulation that will prove as deceptive.

RQ4:

This will be purely exploratory research. We will build up to a model with every variable measuring individual differences and evaluate the results. We will also check correlation between metrics measuring individual differences.

2.4 Results

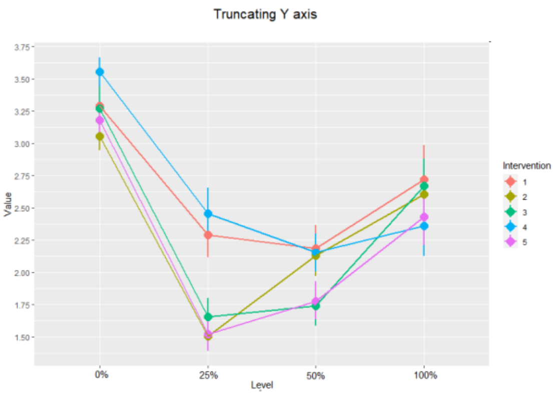
RQ1: *How deceptive are the chosen deceptive techniques? Particularly how is the effect of deceptive techniques related to the level of manipulation.*

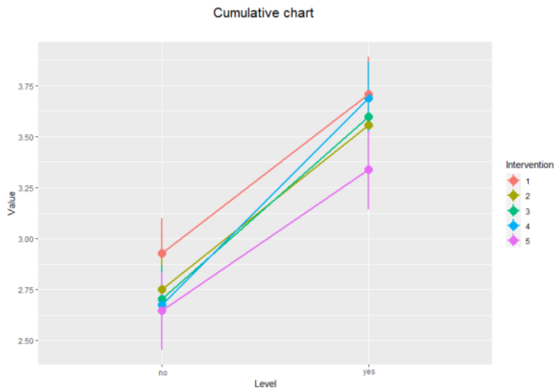
The effect of levels is statistically significant for 7 out of the 9 techniques tested. The only **2** techniques not found to be deceptive are pie chart vs bar chart and pie chart-sorting. But while not significant for all levels, post hoc analysis shows, that there is a significant difference (adjusted p (Tukey) = 0.049) in Pie chart vs. bar charts between levels 3 and 4 (3D pie and bar chart), which as we can see in **Figure 13** is caused by lower average answer in intervention group 2.

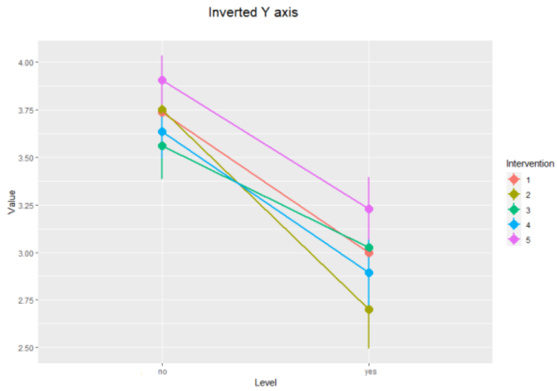
Similar situation is in Quantity as area/ diameter technique, where one intervention groups answer caused a statistically significant difference but according to the ANCOVA, the effect of level here is significant, $F(1, 719) = 6.49$, $p = .011$, ges = 0.0013. Following table shows the statistics for all significant manipulations sorted by generalized eta squared.

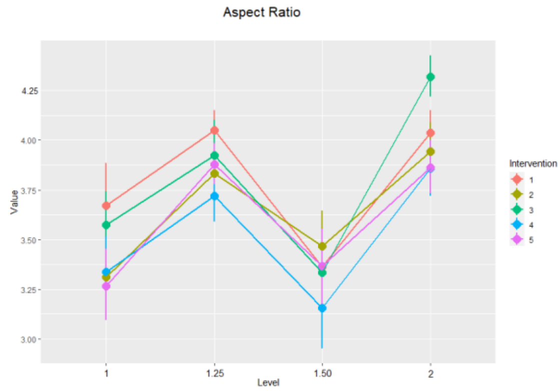
Manipulation	Effect	DFn	DFd	F	p	ges	p adjusted (Holm)
Truncating Y axis	Level	3	2157	306.6688	<0.001	0.2061	<0.001
Cumulative graph	Level	1	719	362.4455	<0.001	0.1281	<0.001
Inverted Y axis	Level	1	719	142.2128	<0.001	0.102	<0.001
Aspect ratio	Level	3	2157	141.3572	<0.001	0.0748	<0.001
Pie chart- rotation	Level	2	1438	7.1464	0.0008	0.0037	0.002
Logarithmic graph	Level	2	1438	4.6848	0.0094	0.0026	0.019
Quantity as area/ diameter	Level	1	719	6.4888	0.0111	0.0013	0.022
Pie chart vs. Bar charts	Level	3	2157	877345	0.064	0.0016	0.127
Pie chart- sorting	Level	1	719	0.0086	0.926	<0.001	0.926

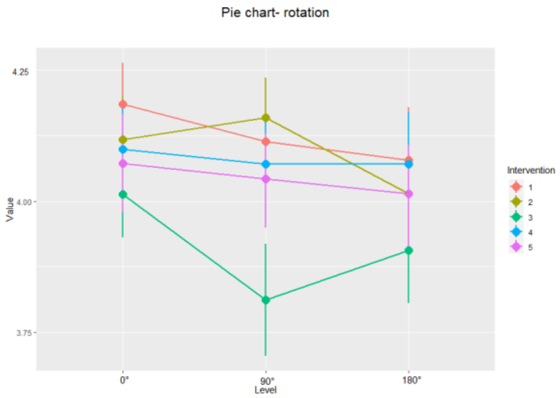
Figure 13- The results for effect of levels reported for each intervention

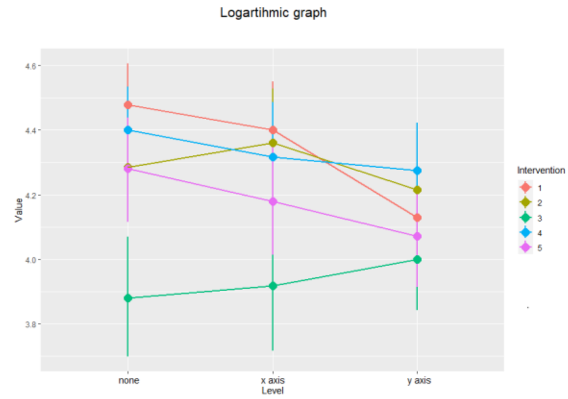

Truncating Y axis
Value
Level
0%
25%
50%
100%
Intervention
1
2
3
4
5

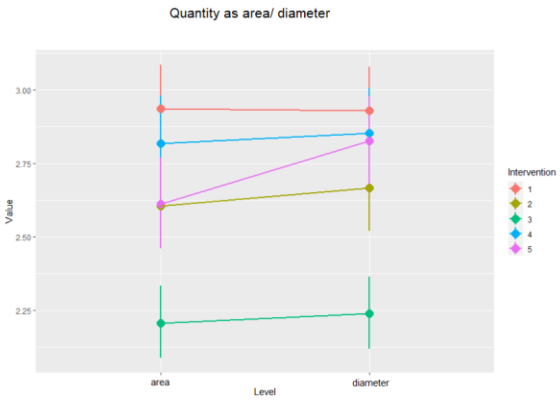

Cumulative chart
Value
Level
no
yes
Intervention
1
2
3
4
5


Inverted Y axis
Value
Level
no
yes
Intervention
1
2
3
4
5


Aspect Ratio
Value
Level
1
1.25
1.50
2
Intervention
1
2
3
4
5


Pie chart- rotation
Value
Level
0°
90°
180°
Intervention
1
2
3
4
5


Logartihmic graph
Value
Level
none
x axis
y axis
Intervention
1
2
3
4
5


Quantity as area/ diameter
Value
Level
area
diameter
Intervention
1
2
3
4
5

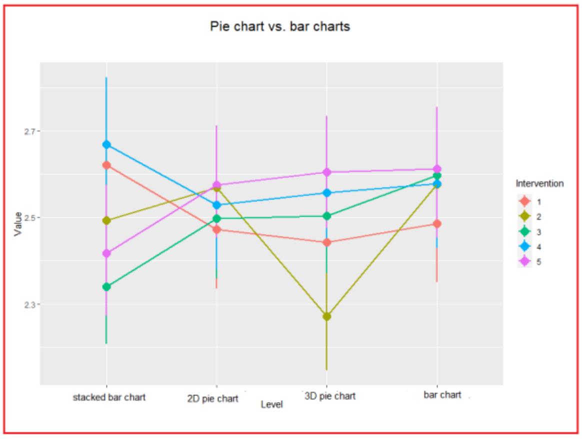
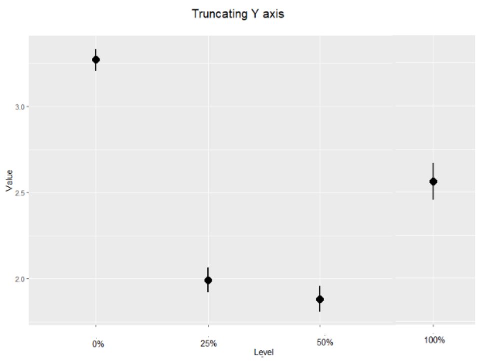

Pie chart vs. bar charts
Value
Level
stacked bar chart
2D pie chart
3D pie chart
bar chart
Intervention
1
2
3
4
5

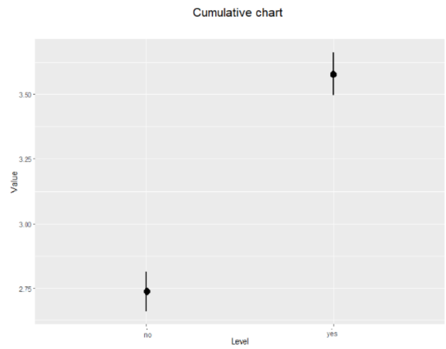
Figure 14- The effect on level for each manipulation split by intervention group. The manipulation in red was showed no significant difference between any levels

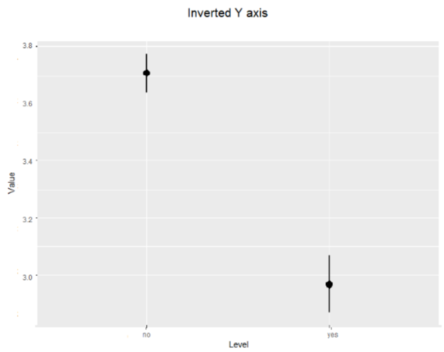
We also calculated estimated marginal means for each pair of levels in each manipulation, which can be found in **Figure 15**. Complete EMM with interaction between level and intervention are for their size uploaded to osf.io/k59zq/. For quick summary we visualized the sizes of EMM for each manipulation on **Figure 16**, but the results need to be contrasted against **Figure 14** to get the whole picture.

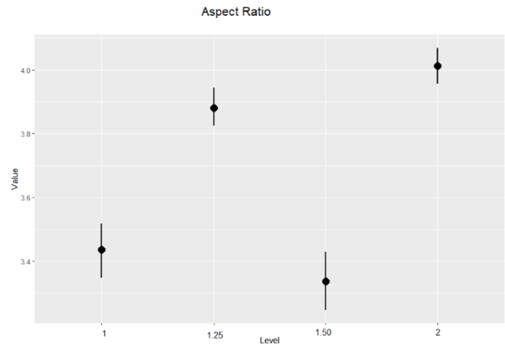
contrast	estimate	SE	df	t.ratio	p value
pie chart rotation 180 - pie chart rotation 90	-0.0232	0.0224	1438	-1.034	0.555
pie chart rotation 0 - pie chart rotation 90	**0.0576**	**0.0224**	**1438**	**2.5673**	**0.028**
pie chart rotation 0 - pie chart rotation 180	**0.0808**	**0.0224**	**1438**	**3.6013**	**0.001**
logarithmic x - logarithmic y	**0.0971**	**0.0412**	**1438**	**2.3592**	**0.048**
logarithmic none - logarithmic y	**0.1271**	**0.0412**	**1438**	**3.0863**	**0.006**
logarithmic none - logarithmic x	0.0299	0.0412	1438	0.7271	0.747
truncating Y axis25 - truncating Y axis50	-0.1124	0.0515	2157	-2.1826	0.128
truncating Y axis100 - truncating Y axis50	**0.5595**	**0.0515**	**2157**	**10.8634**	**<0.001**
truncating Y axis100 - truncating Y axis25	**0.6719**	**0.0515**	**2157**	**13.046**	**<0.001**
truncating Y axis0 - truncating Y axis50	**302057**	**0.0515**	**2157**	**24.7091**	**<0.001**
truncating Y axis0 - truncating Y axis25	**712590**	**0.0515**	**2157**	**26.8917**	**<0.001**
truncating Y axis0 - truncating Y axis100	**0.7131**	**0.0515**	**2157**	**13.8457**	**<0.001**
aspect ratio 1.50 - aspect ratio 2	**-0.6677**	**0.0394**	**2157**	**-16.96**	**<0.001**
aspect ratio 1.25 - aspect ratio 2	**-0.1239**	**0.0394**	**2157**	**-3.147**	**0.009**
aspect ratio 1.25 - aspect ratio 1.50	**0.5438**	**0.0394**	**2157**	**13.8129**	**<0.001**
aspect ratio 1 - aspect ratio 2	**-0.5726**	**0.0394**	**2157**	**-14.5433**	**<0.001**
aspect ratio 1 - aspect ratio 1.50	0.0951	0.0394	2157	2.4167	0.074
aspect ratio 1 - aspect ratio 1.25	**-0.4487**	**0.0394**	**2157**	**-11.3962**	**<0.001**
inverted no - inverted yes	**0.7469**	**0.0624**	**719**	**11.9724**	**<0.001**
quantity as area - quantity as radius	**-0.0676**	**0.0261**	**719**	**-2.593**	**0.01**
cumulative no - cumulative yes	**-0.8364**	**0.0441**	**719**	**-18.9697**	**<0.001**

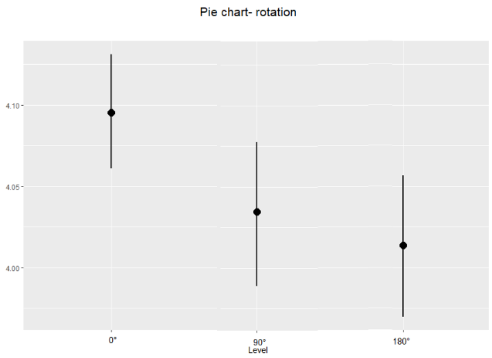
Figure 15- EMM pairwise comparison of levels

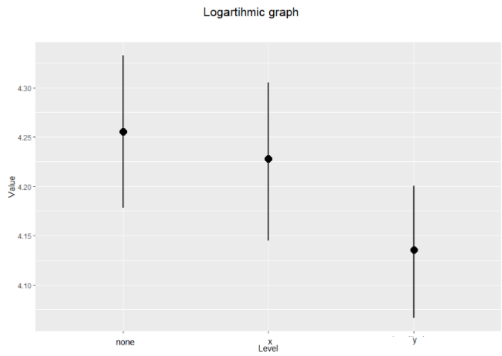

Truncating Y axis
Value
3.0
2.5
2.0
0%
25%
50%
100%
Level

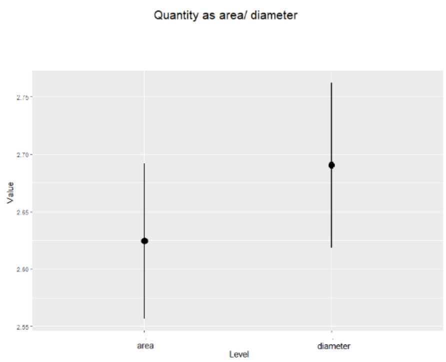

Cumulative chart
Value
3.50
3.25
3.00
2.75
no
yes
Level

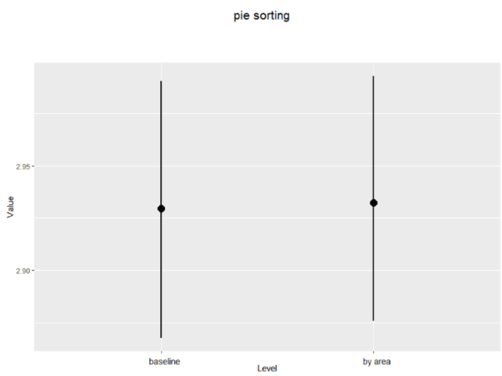

Inverted Y axis
Value
3.8
3.6
3.4
3.2
3.0
no
yes
Level


Aspect Ratio
Value
4.0
3.8
3.6
3.4
1
1.25
1.50
2
Level


Pie chart- rotation
4.10
4.05
4.00
0°
90°
180°
Level


Logartihmic graph
Value
4.30
4.25
4.20
4.15
4.10
none
x
y
Level


Quantity as area/ diameter
Value
area
diameter
Level


pie sorting
Value
2.95
2.90
baseline
by area
Level

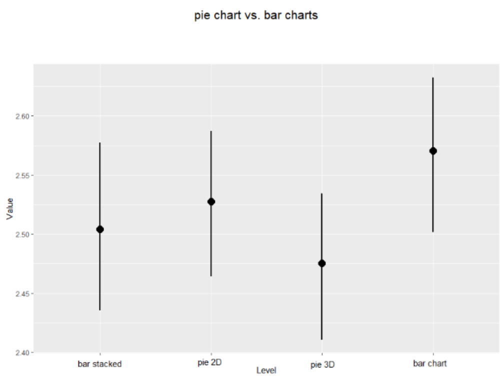

pie chart vs. bar charts
Value
2.60
2.55
2.50
2.45
2.40
bar stacked
pie 2D
pie 3D
bar chart
Level

Figure 16 the differences in means for each graph without splitting by intervention group. Results need to be compared with Figure 14 as not to miss effect of intervention which is not visible here.

RQ2: *Does providing more detailed warning about deceptive visualizations lower their effect on viewer?*

We found no conclusive proof that providing users with more information about deceptive visualization has any effect on user, which would span across multiple deceptive techniques, however there are some statistically significant differences for specific manipulations. Their benefit for practical use is discussed in Discussion

The strongest effect we found ($F(4, 719) = 15.84$, $p < 0.001$, ges = 0.07) was on Quantity as area/ diameter manipulation, where group in intervention group 5 had a significant difference in answers on each level of manipulation (Mean difference = 0.0698, 95% CI = [0.0778, 0.354]), meaning this group on average answered 0.0698 points higher when rating effect shown as diameter instead of as area.

This is the biggest effect of intervention found for all interventions combined, but for us the differences between each intervention group and the control group are more interesting, as reported by the EMM. The table with complete set EMM for all pairs can again be found at osf.io/k59zq/.

Interesting is the difference between intervention 1 and 3 for manipulations Quantity as area/ diameter, Logarithmic chart, Pie chart rotation, and Truncating Y axis. For each of these tactics the estimated difference of means is significant, and the estimate is not trivial: for comparison between Level 1 and 2 for the first two techniques mentioned it is 0.71 for both with SE= 0.098. As we can see in the figure XY users in intervention group 3 were on average more conservative with their answers and usually rated the effect seen as lesser then the rest of the groups.

RQ3: *How does the level of visual literacy relate to the extent to which a viewer is deceived by deceptive tactics?*

Correlation Matrix

Correlation Matrix

		Truncating Y axis 25%	Score	Truncating Y axis 50%	Truncating Y axis 100%	Inverting Y axis yes	Aspect ratio 2	Aspect ratio 1.25
Truncating Y axis 25%	Pearson's r	—						
	p-value	—						
	Spearman's rho	—						
	p-value	—						
Score	Pearson's r	−0.124 ***	—					
	p-value	< .001	—					
	Spearman's rho	−0.123 ***	—					
	p-value	< .001	—					
Truncating Y axis 50%	Pearson's r	0.717 ***	−0.077 *	—				
	p-value	< .001	0.037	—				
	Spearman's rho	0.700 ***	−0.109 **	—				
	p-value	< .001	0.003	—				
Truncating Y axis 100%	Pearson's r	0.531 ***	−0.222 ***	0.543 ***	—			
	p-value	< .001	< .001	< .001	—			
	Spearman's rho	0.519 ***	−0.208 ***	0.526 ***	—			
	p-value	< .001	< .001	< .001	—			
Inverting Y axis yes	Pearson's r	−0.033	0.169 ***	−0.090 *	−0.445 ***	—		
	p-value	0.381	< .001	0.016	< .001	—		
	Spearman's rho	−0.069	0.167 ***	−0.106 **	−0.440 ***	—		
	p-value	0.065	< .001	0.004	< .001	—		
Aspect ratio 2	Pearson's r	−0.219 ***	0.102 **	−0.199 ***	−0.345 ***	0.142 ***	—	
	p-value	< .001	0.006	< .001	< .001	< .001	—	
	Spearman's rho	−0.224 ***	0.078 *	−0.201 ***	−0.350 ***	0.140 ***	—	
	p-value	< .001	0.036	< .001	< .001	< .001	—	
Aspect ratio 1.25	Pearson's r	−0.299 ***	0.118 **	−0.273 ***	−0.445 ***	0.166 ***	0.765 ***	—
	p-value	< .001	0.002	< .001	< .001	< .001	< .001	—
	Spearman's rho	−0.284 ***	0.106 **	−0.259 ***	−0.449 ***	0.188 ***	0.767 ***	—
	p-value	< .001	0.004	< .001	< .001	< .001	< .001	—

Note. * p < .05, ** p < .01, *** p < .001

Figure 17 Correlation matrix for the statistically significant correlations

We found significant correlation with Deceptive delta (difference between a particular level of deception and the non-deceptive version) in 6 levels spanning 3 different manipulations. The correlation matrix with significant values is in **Figure 187**. The 2 cases with highest Spearman's rho were Truncating Y axis 100% with rho = -0.222 and Inverted Y axis (Spearman's rho = 0.169). To visualize the interaction effect, we plotted score on x axis, value on y axis and split the chart by intervention group. We created that for each manipulation, result of which can be found on **Figure 18**.

RQ4: *Is there any relationship between individual differences and the effectiveness of deceptive techniques?*

We used mixed model ANOVA and in addition calculated correlation of collected information with DD, but found no effect with practical significance, nor any, that would make sense based on theory. The analysis done are uploaded to OSF. A discovery of interest is that visual literacy score does not correlate with neither education level (p = 0.061, Spearman's rho = 0.073) nor with style of thinking (pictures, words, something in between) p = 0.303, Spearman's rho = -0.046.

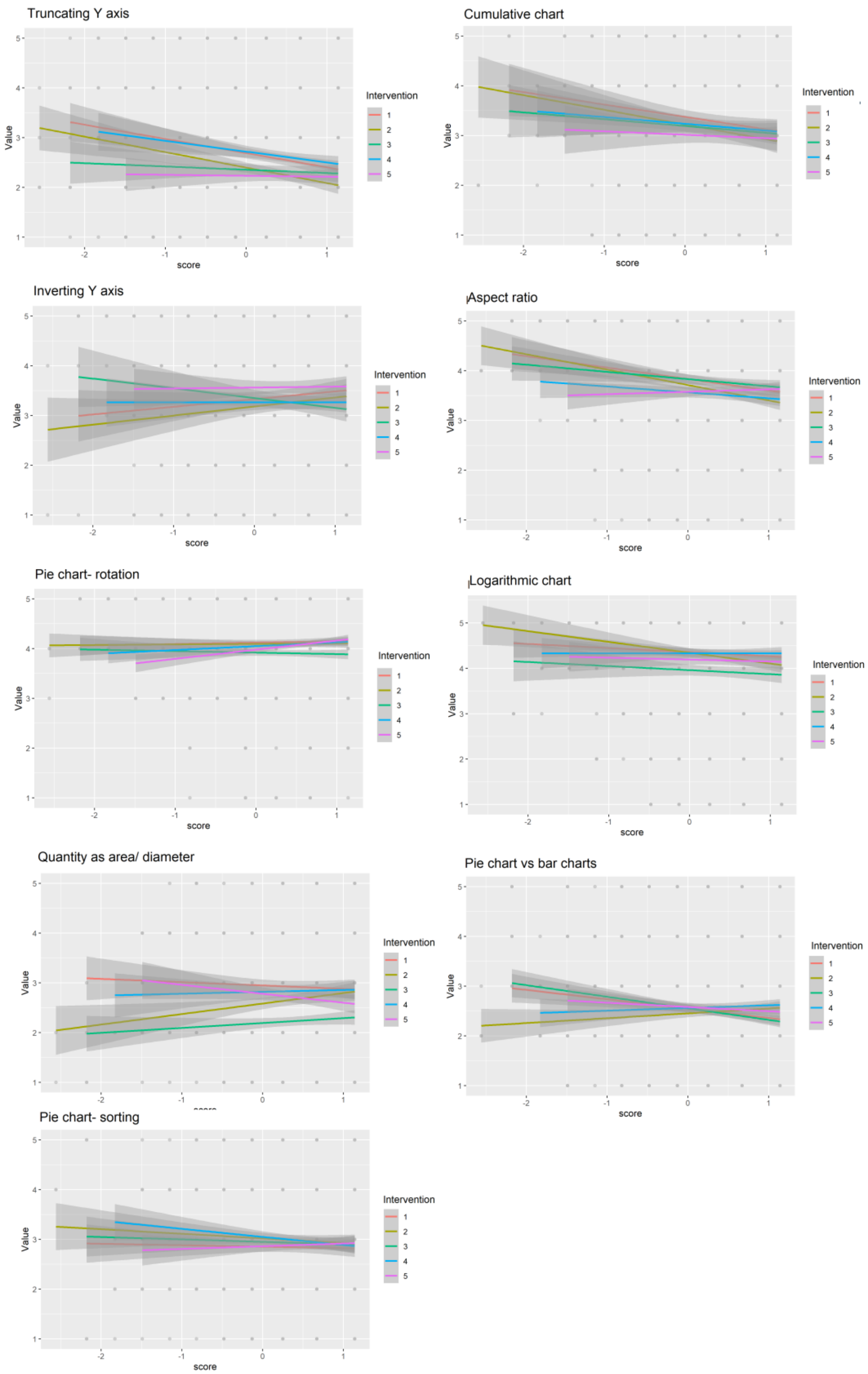


Figure 19- Interaction effect of score and intervention group on the participants' answers

2.5 Discussion

2.5.1 Assessment of results

We measured the effectiveness of deceptive data visualizations, confirming the findings of Pandey et al. (2015) on all manipulations except for Quantity as area/ diameter, there we observed no significant effect, which is interesting, because it is believed to be one of the primary deceptive techniques.

We also discovered two new visualization manipulation with the potential to be misleading: rotating pie chart (in our case only the level with 180° rotation was significant) and transforming the axis of a line chart to logarithmic scale (we found significant deceptive effect for transforming Y axis in contrast to control).

We found no evidence that any of our information protected the user against all (or most) of deceptive techniques. Even effect of particular interventions on specific graphs (which can be inflated due to the large number of comparisons made) is very small as compared to the effect Camba et al. (2022) found for their interventions and so we would not recommend it if guaranteed improvement is required, but we do not close this branch of research, as different kinds of textual intervention might show better results.

We also discovered an interesting fact: that deceptiveness of Truncating Y axis and changing aspect ratio does not increase linearly with the level of change as can be seen on **Figure 16.**

In future research it would be interesting to test, how the interventions put forward by Camba influence the effect of deceptive visualizations (as he only measured whether the participant correctly recognize the deceptive tactic as misleading).

2.5.2 Limits of the study

As was mentioned in Design section, there was a difference between planned study design and the used design due to technical issues. As it happened, we could not control for the effect of seeing the stimuli in a particular order. Even though he ges for intervention are classified as small (not noticeable by experts), we cannot be sure that they wouldn't change if we randomized the paths as planned, but we suspect that it would not make a noticeable difference.

Also, we did not control whether the users read the intervention, so some of the users might have not received the intervention, in future research some sort of verification that the user has completely read the intervention should be deployed.

Conclusion

The aim of the theoretical part of the thesis was to summarize relevant aspects related to deceptive visualizations. In the practical part, we asked several research questions based on current research on this topic. We were interested in whether the deceptiveness of deceptive techniques can be reduced by a one-time intervention in the form of a text warning about this phenomenon varying in the levels of detail. We hypothesized that it could, and that the higher the level of intervention, the less effect the techniques would have on users. This assumption was not confirmed. We do not discard this approach, but so far it seems inferior to interventions invented and measured by Camba et al. (2022).

At the same time, we also investigated the effect size of several validated and some unexplored deceptive techniques, discovering that pie chart rotation and line graph displays on logarithmic axes can be misleading.

Our analysis also shows that visual score results as measured by "Line graph 1" test developed by Boy et al. (2014) does not influence the effect, the deceptive visualizations nor interventions will have on participant in a significant way.

References:

Avgerinou, M., & Ericson, J. (2002). A Review of the Concept of Visual Literacy. *British Journal of Educational Technology*, *28*, 280–291. https://doi.org/10.1111/1467-8535.00035

Bond, C. F., & Depaulo, B. M. (2006). Accuracy of Deception Judgments. *Personality and social psychology review*, 10(3), 214-234. https://doi.org/10.1207/s15327957pspr1003_2

Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., & Pfister, H. (2013). What Makes a Visualization Memorable? *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2306–2315. https://doi.org/10.1109/TVCG.2013.234

Cairo, A. (2014). Graphics Lies, Misleading Visuals. *New Challenges for Data Design*, 103-116. https://doi.org/10.1007/978-1-4471-6596-5_5

Camba, J. D., Company, P., & Byrd, V. (2022). Identifying Deception as a Critical Component of Visualization Literacy. *IEEE Computer Graphics and Applications*, *42*(1), 116–122. https://doi.org/10.1109/MCG.2021.3132004

Camba, J. D., Company, P., & Byrd, V. (2022). Identifying Deception as a Critical Component of Visualization Literacy. *IEEE Computer Graphics and Applications*, *42*(1), 116–122. https://doi.org/10.1109/MCG.2021.3132004

Cohen, M. A., Dennett, D. C., & Kanwisher, N. (2016). What is the bandwidth of perceptual experience? Trends in Cognitive Sciences, 20, 324–335. https://doi.org/10.1016/j.tics.2016.03.006

Correll, M., Bertini, E., & Franconeri, S. (2020). Truncating the Y-Axis: Threat or Menace? *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–12. https://doi.org/10.1145/3313831.3376222

Debes, J. (1969). The loom of visual literacy: An overview. *Audiovisual Instruction*, 14(8), 25-27

Dragicevic, P., & Jansen, Y. (2018). Blinded with Science or Informed by Charts? A Replication Study. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 781–790. https://doi.org/10.1109/TVCG.2017.2744298

Elting, L. S., Martin, C. G., Cantor, S. B., & Rubenstein, E. B. (1999). Influence of Data Display Formats on Physician Investigators' Decisions to Stop Clinical Trials: Prospective Trial with Repeated Measures. *BMJ: British Medical Journal*, *318*(7197), 1527–1531.

Few, S. (2012). *Show me the numbers: designing tables and graphs to enlighten* (2nd ed). Analytics Press.

Franconeri, S. L., Padilla, L. M., Shah, P., Zacks, J. M., & Hullman, J. (2021). The Science of Visual Data Communication: What Works. *Psychological Science in the Public Interest*, *22*(3), 110–161. https://doi.org/10.1177/15291006211051956

How Charts Lie: Getting Smarter About Visual Information. (2019). *Publishers Weekly*, *266*(34), 122.

Huang, W. (Ed.). (2014). *Handbook of Human Centric Visualization*. Springer New York. https://doi.org/10.1007/978-1-4614-7485-2

Huff, D. (1954). *How to lie with statistics*. Norton.

Jarvenpaa, S. L. (1988). Empirical Investigation of Tufte's "Lie Factor" with Computer Generated Graphics. *IFAC Proceedings Volumes*, *21*(5), 335-338. https://doi.org/10.1016/S1474-6670(17)53930-0

Kosara, R. (2016). An Empire Built On Sand: Reexamining What We Think We Know About Visualization. *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization - BELIV '16*, 162–168. https://doi.org/10.1145/2993901.2993909

Lauer, C., & O'Brien, S. (2020). How People Are Influenced by Deceptive Tactics in Everyday Charts and Graphs. *IEEE Transactions on Professional Communication*, *63*(4), 327–340. https://doi.org/10.1109/TPC.2020.3032053

Lee, C., Yang, T., Inchoco, G. D., Jones, G. M., & Satyanarayan, A. (2021). Viral Visualizations: How Coronavirus Skeptics Use Orthodox Data Practices to Promote Unorthodox Science Online. Proceedings of the 2021 CHI Conference on Human Factors *in Computing Systems*, 1–18. https://doi.org/10.1145/3411764.3445211

McNutt, A. M., Huang, L., & Koenig, K. (2021). *Visualization for Villainy* (arXiv:2109.06007). arXiv. http://arxiv.org/abs/2109.06007

Merriam-Webster. (n.d.). Deception. In *Merriam-Webster.com dictionary*. Retrieved July 23, 2022, from https://www.merriam-webster.com/dictionary/deception

Pandey, A. V., Manivannan, A., Nov, O., Satterthwaite, M., & Bertini, E. (2014). The Persuasive Power of Data Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 2211–2220. https://doi.org/10.1109/TVCG.2014.2346419

Pandey, A. V., Rall, K., Satterthwaite, M., Nov, O., & Bertini, E. (2015). How Deceptive are Deceptive Visualizations?. *Proceedings of the 33rd Annual ACM Conference on human factors in computing systems*, 1469-1478. https://doi.org/10.1145/2702123.2702608

S. Bateman, R. Mandryk, C. Gutwin, A. Genest, D. McDine, and C. Brooks. Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. ACM Conference on Human Factors in Computing Systems (CHI), pages 2573–2582, 2010.

S. Haroz, R. Kosara, and S. L. Franconeri. ISOTYPE Visualization – Working Memory, Performance, and Engagement with Pictographs. In Proceedings CHI, pages 1191–1200, 2015.

Swoboda, H., Císař, J., & Frisch, R. A. K. (1977). *Moderní statistika*. Svoboda.

Tal, A., & Wansink, B. (2016). Blinded with science: Trivial graphs and formulas increase ad persuasiveness and belief in product efficacy. *Public Understanding of Science*, *25*(1), 117–125. https://doi.org/10.1177/0963662514549688

Tufte, E. R. (1983). *The visual display of quantitative information* (2nd ed). Graphics Press.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information Processing in the Primate Visual System: An Integrated Systems Perspective. *Science*, *255*(5043), 419–423. https://doi.org/10.1126/science.1734518

Marriott, K., Purchase, H., Wybrow, M., & Goncu, C. (2012). Memorability of Visual Features in Network Diagrams. *IEEE transactions on visualization and computer graphics*, *18*(12), 2477-2485. https://doi.org/10.1109/TVCG.2012.245

Table of Figures

Attachment 1. : Intervention messages

Intervention 1- None

Intervention 2- There are deceptive techniques

Grafy slouží v dnešní době jako nevynechatelný druh komunikace. Umožňují lehce shrnout velké množství dat jednoduchým a přehledným způsobem. Cílem grafů je jasně a zřetelně předat řečníkovu zprávu, ale existují i techniky, které můžou předávané informace zkreslit. Tyto klamavé způsoby zobrazení dat mají většinou jeden ze tří cílů: přehnat velikost efektu, podhodnotit ho, nebo úplně otočit.

Tyto techniky mohou být použity cíleně, za účelem oklamat čtenáře, ale i neúmyslnou chybou autora. Prosíme, abyste na toto mysleli, až budete odpovídat na následujících 26 otázek.

Intervention 3- + showcase of deceptive technique (text)

Grafy slouží v dnešní době jako nevynechatelný druh komunikace. Umožňují lehce shrnout velké množství dat jednoduchým a přehledným způsobem. Cílem grafů je jasně a zřetelně předat řečníkovu zprávu, ale existují i techniky, které můžou předávané informace zkreslit. Tyto klamavé způsoby zobrazení dat mají většinou jeden ze tří cílů: přehnat velikost efektu, podhodnotit ho, nebo úplně otočit.

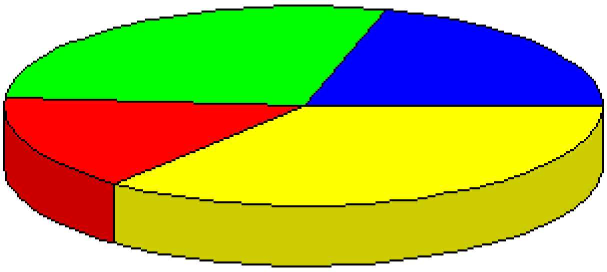
Jednou z nich může být 3D zobrazení koláčového grafu. Změněná perspektiva způsobuje, že se úhly zdají větší či menší, než ve skutečnosti jsou. Výseče v popředí působí, že jsou větší, než ve skutečnosti jsou a naopak s výsečemi v zadní části koláčového grafu.

Tato a jiné techniky můžou být použity cíleně, za účelem oklamat čtenáře, ale i neúmyslnou chybou autora. Prosíme, abyste na toto mysleli, až budete odpovídat na následujících 26 otázek.

Intervention 4- + showcase of deceptive techniques (grafical)

Grafy slouží v dnešní době jako nevynechatelný druh komunikace. Umožňují lehce shrnout velké množství dat jednoduchým a přehledným způsobem. Cílem grafů je jasně a zřetelně předat řečníkovu zprávu, ale existují i techniky, které můžou předávané informace zkreslit. Tyto klamavé způsoby zobrazení dat mají většinou jeden ze tří cílů: přehnat velikost efektu, podhodnotit ho, nebo úplně otočit.

Jednou z nich může být 3D zobrazení koláčového grafu. Změněná perspektiva způsobuje, že se úhly zdají větší či menší, než ve skutečnosti jsou. Například v následující ukázce: žlutá část vypadá větší než zelená, ale je tomu tak? Šikovnou manipulací se takto dají zmenšit nežádoucí výseče a naopak zvýraznit žádoucí.

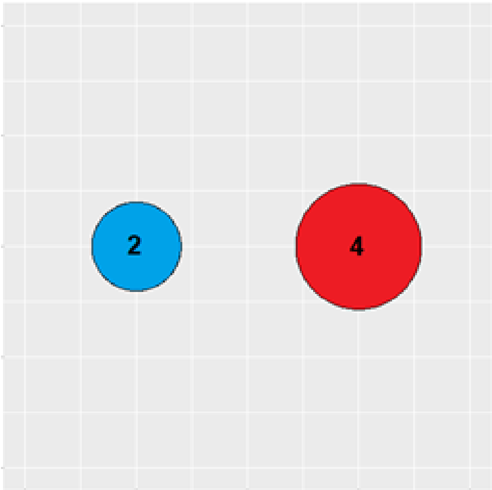


Tato a jiné techniky můžou být použity cíleně, za účelem oklamat čtenáře, ale i neúmyslnou chybou autora. Prosíme, abyste na toto mysleli, až budete odpovídat na následujících 26 otázek.

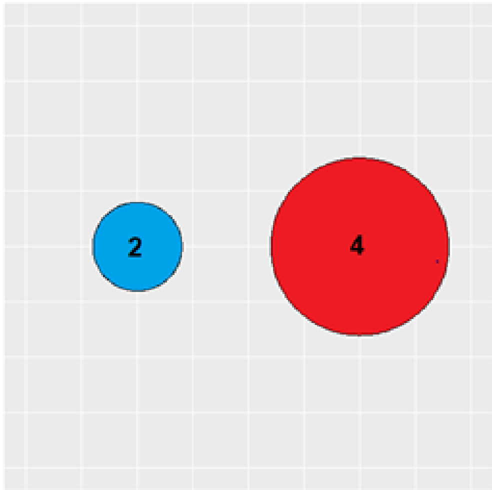
Intervention 5 - + showcase used of deceptive techniques

Grafy slouží v dnešní době jako nevynechatelný druh komunikace. Umožňují lehce shrnout velké množství dat jednoduchým a přehledným způsobem. Cílem grafů je jasně a zřetelně předat řečníkovu zprávu, ale existují i techniky, které můžou předávané informace zkreslit. Tyto klamavé způsoby zobrazení dat mají většinou jeden ze tří cílů: přehnat velikost efektu, podhodnotit ho, nebo úplně otočit.

Jednou z těchto technik je zobrazování velikosti efektu pomocí průměru a plochy.



Velikost jako plocha



Velikost jako průměr

Jak vidíme, změna je v obou případech stejná, ale u zobrazení pomocí průměru působí mnohem výrazněji a při letmém vyhodnocení grafu bychom ji spíš označili za mnohem větší, než ve skutečnosti je.

Tyto techniky můžou být použity cíleně, za účelem oklamat čtenáře, ale i neúmyslnou chybou autora. Prosíme, abyste na toto mysleli, až budete odpovídat na následujících 26 otázek.

Attachment 2: The complete ANCOVA results

1	ffect	Manipulation	DFn	DFd	SSn	SSd	F	p	p<.05	ges	p adjusted (Holm)	Manipulation
2	(Intercept)	Quantity as area/ diameter	1	719	10225.901	1034.9142	7104.3788	0	*	0.8941	0	Quantity as area/ diameter
3	Intervention	Quantity as area/ diameter	4	719	91.185	1034.9142	15.8376	0	*	0.07	0	Quantity as area/ diameter
4	Level	Quantity as area/ diameter	1	719	1.591	176.3111	6.4888	0.0111	*	0.0013	0.022	Quantity as area/ diameter
5	Intervention:Level	Quantity as area/ diameter	4	719	2.098	176.3111	2.1386	0.0744		0.0017	0.074	Quantity as area/ diameter
6	(Intercept)	Inverted Y axis	1	719	16164.533	749.6376	15503.8914	0	*	0.9018	0	Inverted Y axis
7	Intervention	Inverted Y axis	4	719	20.829	749.6376	4.9945	0.0006	*	0.0117	0.001	Inverted Y axis
8	Level	Inverted Y axis	1	719	199.892	1010.6159	142.2128	0	*	0.102	0	Inverted Y axis
9	Intervention:Level	Inverted Y axis	4	719	10.492	1010.6159	1.8661	0.1146		0.0059	0.115	Inverted Y axis
10	(Intercept)	Cumulative graph	1	719	14454.852	1227.7077	8465.4012	0	*	0.893	0	Cumulative graph
11	Intervention	Cumulative graph	4	719	14.941	1227.7077	2.1875	0.0688		0.0086	0.138	Cumulative graph
12	Level	Cumulative graph	1	719	254.454	504.7717	362.4455	0	*	0.1281	0	Cumulative graph
13	Intervention:Level	Cumulative graph	4	719	4.275	504.7717	1.5222	0.1939		0.0025	0.194	Cumulative graph
14	(Intercept)	Truncating Y axis	1	719	17045.814	1328.9961	9221.9532	0	*	0.8339	0	Truncating Y axis
15	Intervention	Truncating Y axis	4	719	79.19	1328.9961	10.7107	0	*	0.0228	0	Truncating Y axis
16	Level	Truncating Y axis	3	2157	881.41	2066.5093	306.6688	0	*	0.2061	0	Truncating Y axis
17	Intervention:Level	Truncating Y axis	12	2157	100.08	2066.5093	8.7052	0	*	0.0286	0	Truncating Y axis
18	(Intercept)	Linegraph aspect ratio	1	719	38930.222	1728.6778	16192.0458	0	*	0.9299	0	Linegraph aspect ratio
19	Intervention	Linegraph aspect ratio	4	719	32.6	1728.6778	3.3898	0.0093	*	0.011	0.009	Linegraph aspect ratio
20	Level	Linegraph aspect ratio	3	2157	237.388	1207.4519	141.3572	0	*	0.0748	0	Linegraph aspect ratio
21	Intervention:Level	Linegraph aspect ratio	12	2157	23.66	1207.4519	3.5222	0	*	0.008	0	Linegraph aspect ratio
22	(Intercept)	Pie chart- rotation	1	719	35588.98	438.5283	58350.7938	0	*	0.9807	0	Pie chart- rotation
23	Intervention	Pie chart- rotation	4	719	13.159	438.5283	5.3936	0.0003	*	0.0185	0.001	Pie chart- rotation
24	Level	Pie chart- rotation	2	1438	2.598	261.3453	7.1464	0.0008	*	0.0037	0.002	Pie chart- rotation
25	Intervention:Level	Pie chart- rotation	8	1438	3.39	261.3453	2.3319	0.0173	*	0.0048	0.017	Pie chart- rotation
26	(Intercept)	Pie chart- sorting	1	719	12438.906	784.165	11405.2194	0	*	0.9245	0	Pie chart- sorting
27	Intervention	Pie chart- sorting	4	719	4.929	784.165	1.1298	0.3412		0.0048	0.682	Pie chart- sorting
28	Level	Pie chart- sorting	1	719	0.003	232.1632	0.0086	0.9263		0	0.926	Pie chart- sorting
29	Intervention:Level	Pie chart- sorting	4	719	1.834	232.1632	1.42	0.2256		0.0018	0.677	Pie chart- sorting
30	(Intercept)	Pie chart vs bar chart	1	719	18381.083	1222.1316	10813.8914	0	*	0.8903	0	Pie chart vs bar chart
31	Intervention	Pie chart vs bar chart	4	719	4.786	1222.1316	0.7038	0.5895		0.0021	0.589	Pie chart vs bar chart
32	Level	Pie chart vs bar chart	3	2157	3.522	1042.0387	2.4302	0.0635		0.0016	0.127	Pie chart vs bar chart
33	Intervention:Level	Pie chart vs bar chart	12	2157	18.439	1042.0387	3.1807	0.0002	*	0.0081	0	Pie chart vs bar chart
34	(Intercept)	Logarithmic chart	1	719	38428.405	1312.4574	21052.1301	0	*	0.946	0	Logarithmic chart
35	Intervention	Logarithmic chart	4	719	52.471	1312.4574	7.1862	0	*	0.0234	0	Logarithmic chart
36	Level	Logarithmic chart	2	1438	5.736	880.2953	4.6848	0.0094	*	0.0026	0.019	Logarithmic chart
37	Intervention:Level	Logarithmic chart	8	1438	10.636	880.2953	2.1717	0.0271	*	0.0048	0.027	Logarithmic chart

	Mauchly's Test for Sphericity: Effect	W	p2	p<0.05	Sphericity Corrections: Effect	GGe	p (GG)	p (GG) <0.05	HFe	p (HF)	p (HF) <0.05
2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
6	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
7	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
8	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
9	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
10	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
11	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
12	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
13	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
14	typ_grafu	0.719	0	*	typ_grafu	0.8397	0	*	0.8429	0	*
15	Intervention:Level	0.719	0	*	Intervention:Level	0.8397	0	*	0.8429	0	*
16	Level	0.719	0	*	Level	0.8397	0	*	0.8429	0	*
17	Intervention:Level	0.719	0	*	Intervention:Level	0.8397	0	*	0.8429	0	*
18	typ_grafu	0.7565	0	*	typ_grafu	0.8606	0	*	0.8639	0	*
19	Intervention:Level	0.7565	0	*	Intervention:Level	0.8606	0.0001	*	0.8639	0	*
20	Level	0.7565	0	*	Level	0.8606	0	*	0.8639	0	*
21	Intervention:Level	0.7565	0	*	Intervention:Level	0.8606	0.0001	*	0.8639	0	*
22	typ_grafu	0.9806	0.0009	*	typ_grafu	0.981	0.0009	*	0.9837	0	*
23	Intervention:Level	0.9806	0.0009	*	Intervention:Level	0.981	0.0181	*	0.9837	0.02	*
24	Level	0.9806	0.0009	*	Level	0.981	0.0009	*	0.9837	0	*
25	Intervention:Level	0.9806	0.0009	*	Intervention:Level	0.981	0.0181	*	0.9837	0.02	*
26	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
27	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
29	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
30	typ_grafu	0.947	0	*	typ_grafu	0.9671	0.0656		0.9715	0.07	
31	Intervention:Level	0.947	0	*	Intervention:Level	0.9671	0.0002	*	0.9715	0	*
32	Level	0.947	0	*	Level	0.9671	0.0656		0.9715	0.07	
33	Intervention:Level	0.947	0	*	Intervention:Level	0.9671	0.0002	*	0.9715	0	*
34	typ_grafu	0.8302	0	*	typ_grafu	0.8549	0.0132	*	0.8567	0.01	*
35	Intervention:Level	0.8302	0	*	Intervention:Level	0.8549	0.0356	*	0.8567	0.04	*
36	Level	0.8302	0	*	Level	0.8549	0.0132	*	0.8567	0.01	*
37	Intervention:Level	0.8302	0	*	Intervention:Level	0.8549	0.0356	*	0.8567	0.04	*