Radboud University

Bachelor thesis 2021

Attention to detail!

The influence of involvement and deception on successful graph understanding

Radboud University Nijmegen, Netherlands

Study degree: International Business Communication (Bachelor)

Supervisor:

Student: Tobias Dehncke

Student number:

Date and place of submission: 17.05.2021

Email:

Telephone:

Abstract

In this study, the influences of deception and involvement on understanding of visual communication using graphs have been explored through an experiment. To that end, a high- and a low-involvement condition were evoked through including a relevant and a non-relevant location in the title, description, and questions regarding a graph. To create a deceptive and a non-deceptive condition, the axis of one graph were distorted while another graph was not manipulated. The data in both graphs remained the same. After analyzing the data, the results showed that the manipulations did not lead to any significant effects. Finally, recommendations for future studies in the same field with improved methodology are made.

Introduction

Nowadays, diagrams are an essential part of daily life. In general, graphical representations are used to complement and increase the density of information being conveyed (Durand, Yen, O'Malley, Elwyn & Mancini, 2020). They are nearly omnipresent, whether online or offline: On social media, on news sites, on television, in print media, or in education (Maichle, 1994). They also fulfill the role of communicating important information (Okan, Galesic & Garcia-Retamero, 2016) in an easy-to-process manner to a wide audience (Okan, Stone & de Bruin, 2018) and can improve understanding, and they can influence decisions (Nayak et al., 2016). This advantage is achieved through the supporting role that allows for non-numeric (Durand et al., 2020) or non-textual processing of information (Nayak et al., 2016). This facilitation allows to reach a wider audience than would be the case with purely textual information. In addition to making information more accessible, graphical representations are often used to support decisions. In the medical field these are decisions of patients regarding treatments, to weigh risks and rewards against one another (Okan, et al., 2016). Thus, graphics are data presentations which make data portrayed more easily accessible to viewers and lead to increasing awareness of different aspects of life in public (Nelson, Reyna, Fagerlin, Lipkus & Peters, 2008.

Overall, the goal of graphical representations is to increase the accessibility of information and reduce aspects such as bias, anecdotal reasoning, or framing effects (Durand et al., 2020). One of the most common formats are bar charts (Durand et al., 2020; Peebles & Ali, 2009).

The benefits of visual communication

Stone, Reeder, Parillo, Long, and Walb (2018) compared the influence of depicting information numerically to depicting the information graphically. When presenting information in a written format but making it more prominent (bold and red) did not have any effects. Graphical display of the same data did have the benefit of increased risk awareness for viewers. Thus, graphically displaying information increases awareness and perception of information presented compared to written only information Stone et al. (2018).

The same advantage of visual communication compared to non-visual communication was shown by Garcia-Retamero and Galesic (2010). In their study, they did research on people with difficulties understanding purely numerically displayed data. They showed that visual communication increased the accuracy of participants' perception of risks and risk reduction from 36% to 61% when depicting absolute numerical information. The same effect was achieved when depicting relative numerical information visually, that is an increase in accuracy from 33% to 50% (Gracia-Retamero & Galesic, 2010).

However, visualizations of data can be tricky. Peck, Ayuso, and El-Etr (2019) found different confusing features which reduce the clarity of graphic representations of data. Charts which convey a lot of information, which use differently dotted lines or color schemes were perceived as rather confusing by participants. Generally, people prefer simple charts which are easily understood (Peck et al., 2019).

In addition to unclear graphics, visualizations available to the public often contain misleading features. Such features can be improperly scaled axes or longer bars representing lower values (Okan et al., 2016). In a sample of 74 charts in pharmaceutical advertisements in medical journals, nearly 40% contained misleading features (Okan et al., 2016).

Given the constant presence of graphs in today's information society, the influence of deceptive features in graphs is an important topic to study. How exactly these features influence viewers. How they are understood. And most importantly, if the goal of making information accessible is still achieved or not when such features are present. For a variety of reasons, this goal might not be achieved: Insufficient capabilities of viewers to process a graph correctly, or insufficient attention leading to misinterpretations, or to the inability to identify misleading features can prevent correct comprehension.

Graph processing

According to Roser (1990), the ability individuals possess regarding graph analysis appears to have a higher impact than motivation for graph processing. Overall, perceived relevance, thus attention and motivation for more effortful processing, are intertwined and often correlate. Involvement is therefore a key variable regarding the consequences of information intake (Roser, 1990).

How attention comes into play in processing visualizations can be seen in the work of Padilla, Creem-Regehr, Hegarty, and Stefanucci (2018). According to their research on cross-disciplinary frameworks, they describe the process of interpreting visualizations as a multi-staged process. The process itself is a process which consists of two possible routes to process data which is viewed further.

On the one hand there is low-effort processing, that is the exercise of low amounts of working memory to process the information further (Padilla et al., 2018). This type of processing is used for simpler decisions, which can follow certain rules of thumb and does not require much mental transformation. An example for the processing would be reading a certain number off a chart.

On the other hand, processing which uses more effort, or more working memory, is used to consider the viewed information more and process it deeper (Padilla et al., 2018). This process is slower but more considered. In that sense, the high working memory processing is reserved for more complex computations, so if the mental presentation needs to be transformed to match a certain mental schema. An example for information which is processed in that way could be calculating the average between two data-points in a chart.

The process of comprehension and interpretation itself, as described by Padilla et al. (2018), is divided into several smaller steps. First, the visual input is shown to a viewer. Next, this physical image is transformed into a mental representation of it. Based on this mental image, information stored in the long-term memory of viewers which fits the visual information is searched. Information which suits the presented graph is also called a “mental schema” (Padilla et al., 2018). Any other information stored in the long-term memory, which might be helpful, is searched for in this stage. After both, the mental schema and further information are retrieved, a matching process between the mental representation of the information and the mental schema are takes place. The information surrounding the graph can also help with the graph interpretation process. The interpretation itself is the final product of the visualization process, which entails the data which is the message that was encoded in the visual image. This final product is called “conceptual message” by Padilla et al (2018).

This final product of interpretation entails any mental influences which a viewer had: Additional information used to interpret, mental changes of the information, or the matching process to a visual schema. The task given to viewers may influence the outcome as well as it changes the mental processes required for the information interpretation – the "conceptual message" (Padilla et al., 2018) which is encoded in the task given, requires a matching "conceptual message".

The influence of involvement

On top of the importance of conceptual questions on information processing, attention towards graphs is directed by viewers individually. Peck et al. (2019) tested how people rank the perceived importance of graphs, and to what extent participants then direct attention towards a specific graph. In their study, they found out that locality of information has a strong influence on attention and involvement of participants, similar to the type of information. If the information presented was perceived as having a personal link to viewers, the information was evaluated as more relevant. An example would be information about drug use if viewers had a personal connection to substance abuse. This personal link was also the case if the location of a graph showed a location where the viewer lives or lived before. Any graph which had some indication of locality, be it national or local, showed the same trends. This indication of locality also had implications on participants' evaluations of the quality of information depicted as more relevant information was also evaluated as of a higher quality. The attention directed to a graph coincided with graphs perceived as of a higher quality (Peck et al., 2019).

Higher involvement due to personal relevance leads to increased motivation to ensure that a taken decision is the right ones (Liberman & Eagly, 1989). As a result, arguments which are perceived as more personally relevant are considered more carefully and viewed more critically (Johnson & Eagly, 1989). Thus, when participants are more involved, they are likely to pay more attention to the data presented, as also demonstrated by Roser (1990).

To compare high and low involvement conditions, Roser (1990) performed a 2X2 study in groups of high and low motivation to process information. This was done by telling the participants that the information is either irrelevant to them or that the information is going to be useful for a follow-up task. Deep or shallow processing conditions were created by allowing participants to take notes or introducing distracting elements. If the message was perceived as helpful in an essay to be written later, deeper processing took place due to the higher involvement of participants. This increased attention to the message increased retention and led to increased perceived relevance of the message. (Roser, 1990).

This result is consistent with the findings of Petty and Cacioppo (1984), who found that higher involvement leads to deeper processing of the arguments presented. In the case of strong arguments this led to viewers being more convinced. Weaker arguments led to lower persuasive success. Participants were less involved, a higher number of different arguments was more convincing than the quality of arguments (Petty and Cacioppo, 1984). The above-mentioned findings show that attention to information does have influence on graph processing and graph understanding when it is evoked by features embedded in the information display.

On graphicacy

On top of attention directed towards information, the abilities to process graphs of viewers play a role in graph comprehension. Planinic, Ivanjek, Susac, and Milin-Sipus (2013) found that the ability to correctly interpret graphs is also highly context dependent: People are unable to interpret the results of a graph if they do not have sufficient knowledge of the context in which the information is presented. Planinic et al. (2013) tested students, who were able to correctly interpret a graph presented in a mathematics context. Other groups of students were unable to interpret data correctly when the surrounding context was presented to be in the domain of physics, or from a third, physics unrelated field. This suggests that an audience requires some basic skillset in the domain of the data presented to be able to fully process and correctly interpret said graph (Planinic et al., 2013). This is in line with the graph processing framework which is proposed by Padilla et al. (2018). In addition, a basic level of mathematical knowledge is required to correctly interpret a given graph (Planinic et al., 2013).

The ability to process and understand graphs is called graphicacy. Graphicacy is a complex interaction between multiple processes and abilities, such as identifying visual features, translating them into information, and inferring interactions between this information (Okan et al., 2016). In other words, graphicacy is the ability to process, understand, and present information in graphic form - this can be diagrams, tables, graphs, maps, or sketches (Brown et al., 2011).

Individuals with high graphicacy ability can identify and focus on relevant information (Okan et al., 2016), so it is a core concept for successful communication through graphic displays (Brown et al., 2011). Maichle (1994) referred to graphicacy as a "graph schema" that contains important conceptual information, which is the same term as the one used by Padilla et al. (2018). This routine allows skilled individuals to easily access, identify, and focus attention on significant features in any graphic (Maichle, 1994; Okan et al., 2016). This scheme may be different for different graphic formats (Maichle, 1994).

These mental conceptualizations of graphs and the routine nature of their exploration means that they are not the same for everyone (Okan et al., 2016). A higher graphicacy ability also means more focus on relevant aspects of the presented graph that also leads to more accurate graph interpretations. For example, people with higher graphicacy ability can recognize conflicting information and potentially see through deceptive representations (Okan et al., 2016). This might help them to process the graphically presented information more accurately than other viewers with lower graphicacy abilities, not only in general but especially in the case of deceptive features being present.

Viewers with lower graphicacy ability also face other issues. So can a viewer's limited ability to transform the data presented into correct conclusions may lead to misunderstanding of graphically presented information (Durand et al., 2020). This is an issue which many adults face, as they often have limited ability to process and understand graphs. However, this issue might often fly under the radar because many reports made on the skills related to graph understanding show positive developments and increasing graph understanding abilities (Durand et al., 2020).

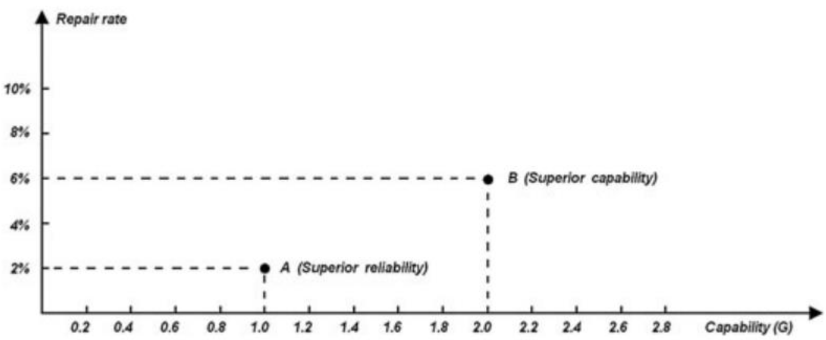
Diagrams are not a panacea for the problem of communicating information clearly and without misunderstandings. The information that is communicated allows people who are used to graphical formats to quickly recognize and infer information from the data presented (Planinic, Ivanjek, Susac & Milin-Sipus, 2013). However, this is not universal. Even though some features of graphs can be identified by novices (Okan, et al., 2016), according to Galesic and Garcia-Retamero (2011), there is no innate humane ability of understanding graphs. Even the simplest graphs can be misunderstood (Galesic & Garcia-Retamero, 2011) and graph misunderstanding is also dependent on different graph formats, depending on which of their main aspects they draw attention to (Peebles & Ali, 2009).

The influence of graph layout

The format of charts also affects the ease of interpreting them correctly. As the layout of graphs draws attention to different parts of the information presented, the focus of viewers is also shifted, which results in either more accurate, or less accurate interpretations of the data presented (Peebles & Ali, 2009). For example, line graphs have been found to be more difficult to understand than others, while bar graphs, which draw attention to the axes, result in a higher number of correct interpretations (Peebles & Ali, 2009).

When graphs, which draw attention to some features over others, are coupled with efforts to reduce noise (in order to limit confusion, for example through the features found by Peck et al., 2019), this can lead to problems: If too much information is left out of the overall picture, it can lead to unintentionally misleading designs of graphs which lead to conclusions that may not be supported by the data presented (Spiller, Reinholtz & Maglio, 2020).

The problems with inferring information from graphs do not stop there. The presentation of data be it within or in the messages around the graph can also be problematic, according to the findings of Sun, Li, Bonini, and Su (2012). Graph framing is a problem because it also affects decisions made in consideration of the data represented by graphs. Graph framing effects had influences on visual perception of graphs through manipulating the visual display of a graph while keeping the numerical information identical. When two different products were compared, the size of the difference between a superior and an inferior product differed significantly (Sun et al., 2012). If the distances portrayed in the visual display were higher, the preference for the superior product was stronger, and when the differences distances portrayed in the visual display were less, the preference for the superior product was weaker (Sun et al., 2012).



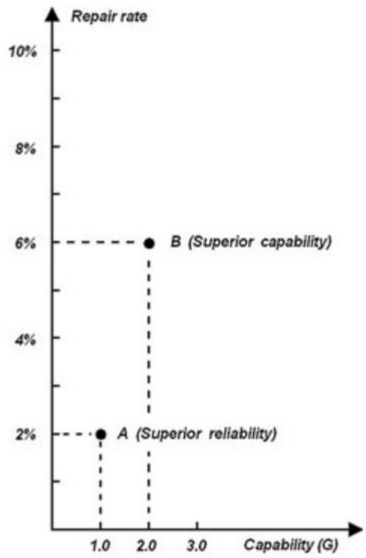


Figure 1 The graphs used by Sun et al. (2012)

Further problems regarding graph interpretation have been found by Spiller et al. (2020). In their study they found that the visualization of data representation matters as the same data viewed in different graphical presentations led to opposite conclusions (Spiller et al., 2020). They compared stock (showing the quantity at any given time) with flow visualizations (showing the change of quantity) of jobs during Obama’s presidency. Even though the numerical data presented to viewers was identical, a difference in evaluation of the data was found. In case of the stock graph the economy was viewed more bullish due to continuously increasing employment which was identifiable due to the visual rise of the graph. In case of the flow chart, viewers perceived the graph as less positive. This may be due to the perceived stagnation of employment growth, even though the number of jobs is still on the rise (Spiller et al., 2020). Different visual presentations of data draw the focus of viewers to different aspects of this data, while also taking the focus away from other aspects of the same data (Peebles & Ali, 2009; Spiller et al., 2020). Thus, the visual presentation of data also has an influence on perception of this data.

In a similar fashion did scaling influence the interpretation of data as has been found by Romano, Sotis, Dominioni, and Guidi (2020). A linear scaling of the y-axis showing Covid-19 cases lead to improved understanding of the data presented, which was rising, while a logarithmic scaling increased misunderstanding and the perception of a “flattening” of the curve (Romano et al., 2020).

Similarly, data can be framed towards individuals who hold different political views, leading to biased conclusions (Peck et al., 2019). This effect of misleading the viewer can already be achieved by graph framing itself (Sun et al., 2012) through a variety of distortion techniques that are common in everyday life (Pandey, Rall, Satterthwaite, Nov & Bertini, 2015). Such an effect can also take place through the message, that is, through the context surrounding the information being presented (Pandey et al., 2015).

Aims of the research

As laid out, previous studies suggest, that there is an influence of involvement on attention on graph perception and understanding. Additionally, deceptive features are an important topic in contemporary research. However, causal relations, proven through an experiment, have not been made established yet. Therefore, the role of the conducted experiment was to assess the influences of involvement and deceptive features on graph understanding. Graphicacy may play a role as mediating factor in this complex process. The exact influences of these features remain to be shown. To verify the influence of these factors, the following research questions were central to the study:

1) What is the influence of involvement on graph understanding?
2) What is the influence of deception on graph understanding?
3) How does involvement interact with deception in regard to graph understanding?

To assess the influences, a between-subject design was used which compares high- and low-involvement conditions to one another. Additionally, deceptive, and regular graphic displays in form of bar graphs were shown to viewers. Further details to the study design and methods follow in the respective sections.

Following the findings of Roser (1990) and Petty and Cacioppo (1984), higher involvement of participants leads to higher attention towards the information portrayed in the graph shown. As this also leads to deeper processing, the higher the involvement is that the participants have, the higher the graph understanding will be. Therefore, also increased understanding of the information will take place.

If participants view the normal bar graph compared to the deceptive bar stock graph, graph understanding is going to be higher than for the deceptive bar stock graph. This hypothesis is based on the findings of Peck et al. (2019), who found a decrease in clarity of graphs which include confusing features, as well as Okan et al. (2016), who found improperly scaled axis to be influential.

A possible interaction effect would play out like this: A high-involvement condition will always lead to higher understanding than a low involvement-condition. This is because more attention is paid to the features of the graphs. A deceptive graph will always lead to lower understanding than a normal bar graph, due to its misleading features.

If a deceptive graph is coupled with high involvement, more attention is paid to the content. This would lead to increased accuracy of the information over the deceptive/low-involvement condition, but less accuracy than the normal bar graph/high involvement condition. If the normal graph is paired up with low-involvement, this would lead to better understanding than the deceptive graph/low-involvement condition, but to worse understanding than the normal graph/high involvement condition.

This leads to the accuracy of condition-ranking as follows: The normal bar stock graph in a high-involvement condition is going to lead to the highest accuracy in the results, both the high-involvement/deceptive bar stock graph and the low-involvement/normal bar stock graph will have results somewhere in between, and the low-involvement/deceptive condition will yield the worst results.

Similarities between the results of the conditions of normal graph/low-involvement and deceptive graph/high involvement were expected. This is due to the low-involvement group being unlikely to pay a lot of attention to the features of the graph, while the high-involvement group makes up for some of the deception due to the increased attention paid to the graph.

Methods

Materials

The experiment consisted of two different graphs shown to participants, and of accompanying messages which were to evoke a high- and a low-involvement condition. Both charts used were bar stock graphs, as this is a familiar chart format, as well as one of the preferred formats in previous studies (Brown et al., 2011; Durand et al., 2020), and information presented is well understood by viewers (Durand et al., 2020). Bar charts lead to high perceptual accuracy (Brown et al., 2011; Garcia-Retamero & Galesic, 2010), limiting potential confounds such as misinterpreting the chart itself, or the information presented, to avoid influences on the outcomes of the study.

Both, the deceptive and non-deceptive graphs were bar stock graphs, indicating the numbers of Covid-19 cases at a given point in time (Spiller et al., 2020). Examples of the graphs are presented in Figure 2, all material used can be found in the appendix.

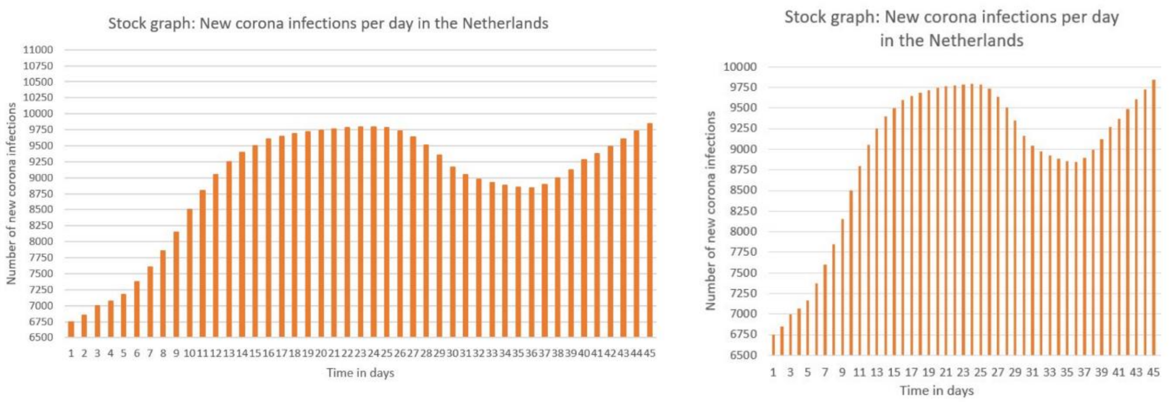


Figure 2 The graphs used to manipulate the deception condition without distortion (left) and with distorted axes (right).

The axes of the deceptive graph were manipulated. To create a distortion effect on the graph (Pandey et al., 2015), the x-axis was compressed to create a less wide image. This resulted in an apparent steeper rising curve than the non-deceptive graph while the information displayed remained the same, similar to the fashion in which the manipulation of the visual display manipulated by Sun et al. (2012) was performed. Furthermore, the maximum of "number of new corona infections" was reduced by 1000, to increase the angle of the curves.

To eliminate noise in the information participants received and to further minimize potential confounds related to the current research, the charts were designed using the techniques suggested by Boers (2018). This is the reduction of unnecessary noise within a graph through minimizing ink, only in core positions (steps of 100 in this case), minimizing the dates by eliminating the year from the date on the axis, and using simple colors that are high in contrast to convey the data clearly (Boers, 2018). Case counts and time trends were presented on the axes only, otherwise manipulation could have been influenced (Garcia-Retamero & Galesic, 2010).

The two accompanying messages which were to elicit high-involvement or low-involvement responses are descriptors of the locations and of what was shown. The high-involvement condition of the Netherlands has been chosen because it is close in proximity and in psychological space, leading to a higher perceived relevance. The condition of Zambia has been chosen due to its spatial and psychological distance to the Netherlands which in turn leads to less perceived relevance, that is lower involvement of participants. The precise messages are presented in Table 1. The Dutch translations of the questions, as well as the complete questionnaire can be found in the appendix.

Table 1 The messages for the high- and low-involvement conditions in English

Condition	Message
Low-involvement condition (English)	The authorities of the Netherlands are supposed to make a decision about whether they should reopen the non-essential shops or whether they should prolong the closing of non-essential shops for another 14 days.
High-involvement condition (English)	On day 45, the authorities of the Zambia are supposed to make a decision about whether they should reopen the non-essential shops or whether they should prolong the closing of non-essential shops for another 14 days.

One of these messages was: “The following stock graph depicts **fictional** data about the number of new corona infections per day in the Netherlands.” In this case the message would elicit a high-involvement response. The message for the low-involvement response is: “The following stock graph depicts **fictional** data about the number of new corona infections per day in Zambia”.

Additionally, the fictionality of the data was also explained in the information and consent form before the study with the sentence: "You are invited to participate in a research project in which you will be presented with a fictional graph about health communication."

This was followed by the question directed to participants, in which they should choose to reopen the local economy or to remain closed by asking: "Based on the graph you are seeing, what would your advice be?". This choice was done using a 7-point Likert-scale ranging from "Definitely stay closed" to "Definitely reopen".

This was done to create a choice for the participants; however, the real choice was not made by them. If participants had to choose for themselves, without the "authorities" as predetermined decision makers, the question itself could lead to higher participation and thus higher perceived relevance for participants.

All questions were asked in versions for the high and the low-involvement conditions, adapted to the graph which was being shown, and an English and a Dutch version of all of the questions of the questionnaire was produced. A complete version of the questionnaire can be found in the appendix.

Participants

The participants of the study were selected by convenience sampling within the social circles of the research group. In order to achieve high and low involvement in the two different message conditions, all participants were recruited in the Netherlands, i.e., either the participants were Dutch, or they had their center of life in the Netherlands at the time of the study.

A total of 391 people participated in the study. Due to the fact that the study was performed in a research group which members had different aims for their own studies, participants in other conditions were excluded. Participants who answered they were younger than 18 in the age check, as well as participants who did not give their consent, or participants who resided in other countries than the Netherlands were sent to the end of the questionnaire and excluded from further analysis. All participants with any missing values were excluded. This left 170 participants to analyze.

Of these 170 participants, 63 were male (37.1%), 105 (61.8%) were female, 1 identified as a different gender, and 1 preferred not to say. The age of participants ranged from 18 to 66 years. The average age was 28.05 ($SE = 12.99$) years.

Of the 170 participants included in the analysis, 81 (47.6%) participants were shown the high involvement condition, while 89 (52.3%) participants were shown the low involvement condition. 86 (50.6%) participants were shown the stock graph, while 84 (49.4%) participants were shown the deceptive version of the graph. 141 (82.9%) participants viewed the Dutch versions of the questionnaire, 29 (17.1%) participants viewed the English version.

Design

The study consisted of a between-subjects design, split into 4 different conditions: standard stock bar chart and deceptive stock bar chart, split into a high-involvement condition and a low-involvement condition, resulting in a total of 4 conditions with 2 independent variables.

In total, 89 participants were shown the low-involvement condition and 81 participants were shown the high involvement condition. 84 participants viewed the deceptive graph, and 86 participants viewed the normal stock graph. The distribution of participants into the 4 conditions can be seen in the table below. All participants were assigned to the groups randomly.

Table 2 The number of participants in each condition

	Low-involvement condition	High-involvement condition	Total
Normal stock graph	46 (27.1%)	40 (23.5%)	86 (50.6%)
Deceptive stock graph	43 (25.3%)	41 (24.1%)	84 (49.4%)
Total	89 (52.4%)	81 (47.6%)	170 (100%)

Instruments

To assess understanding of the graphs shown to participants, five items were asked concerning the graphs and the understanding of information presented. The questions asked were adapted questions based on Okan et al. (2018). The questions also assess inferences made based on the graphs shown. The first question concerned information of the number of Covid-19 cases at one point in time in either the Netherlands or in Zambia, depending on the graph shown: “How many new Covid cases were registered on day 3?”

The next question asked about the increase of cases at a certain time. It was a multiple-choice question: “Look at day 3 and day 4. Which day shows the biggest increase in new infections compared to the day before?" The answer possibilities were either day 3 or day 4. Next, a question regarding the understanding of general trends was asked: "Compare the period from day 1 to day 10 to the period from day 11 to day 20. Which period shows a stronger rise in infections?" This was also a multiple-choice question with the periods named in the question as answer possibilities. The following two questions were open questions, one regarding the difference in new infections on day 12 versus day 13, and the last question was "On which day did the decrease of covid-19 infections start to slow down?". All these questions had both an English, and a Dutch version. For the full questionnaire, consult the appendix. Correct responses were considered as follows: For question 1 it was "7000" cases, for question 2 it was "day 3", for question 3 it was the term from "day 1 to day 10". For question 4, values between 160 and 200 were coded as correct due to the difficulty in reading off the chart. For question 5, the correct response was day "31". After coding the responses into correct and incorrect answers, the reliability of the questionnaire was insufficient (α = .34).

Graphicacy is also to be collected and measured as a control variable. This is done all using the Graph Literacy Scale which Galesic and Garcia-Retamero (2011) developed. It has also been used by Nayak et al. (2016) in previous research. In previous studies this scale had a high inter-item reliability across 13 items measuring graph reading and interpretation skills (α > 0.70) **(**Galesic & Garcia-Retamero, 2011). For the current study this scale has been reduced to 5 items. These items were: graphicacy question 1 – knowing parts of a pie in percent (Q3), graphicacy 2 question – reading off a point on a line chart (Q5), graphicacy question 3 – comparing slopes of a line at 2 intervals (Q6), graphicacy question 4 – reading off a point on a bar chart (Q1), and graphicacy 5 question – assess the difference between 2 bars (Q2) (number of the questions in the original scale in brackets; Galesic & Garcia-Retamero, 2011). The observed reliability of this scale was insufficient after coding the responses into correct and incorrect answers (α = 0.31).

The forced choice was only coded in opening or closing to assess possible influences of the independent variables.

To assess if the manipulation of viewer's attention was successful, manipulation checks were performed. To this end, three 7-point Likert-scales were presented to participants which were based on the work of Frewer, Howard, Hedderley, and Shepherd (1997). The three items were adapted to match the Covid-theme of the overall research. Participants were asked to evaluate to what extent they agreed with the statements presented, ranging from "Strongly disagree" to "Strongly agree". The first item concerned the overall relevance of the graph: "The graph I saw is very relevant to me personally”. The second and the third item concerned the relevance of the topic and regional importance to participants: “The positive COVID-19 tests in the Netherlands are very relevant to me personally”, and “The COVID-19 regulations in the Netherlands are very relevant to me personally”. Statement two and three were adapted to “Zambia” in relevant conditions. As with the other items in the questionnaire, these questions were translated to Dutch. The Dutch translations and the complete questionnaire can be found in the appendix. The reliability of the manipulation checks was good ($\alpha = 0.80$).

Procedure

Before the task, participants were given information on the task and on the procedure, excluding the goal of the study as this might influence the outcomes of the study. A consent form was given to them with contact information regarding the study and of their responsibilities, confirmation that participants can stop at any time, as well as a confirmation form of their age as above the legal age.

As first step of the study, the participants were asked to fill some personal information in a questionnaire as control variables. These consisted of age, gender, highest level of educational achievement, country of origin, and country of residence. These have been asked to assess other potential influences on the outcomes after the study.

In a next step, the participants were shown the message which precedes the graph. This is followed by viewing the message which precedes the graph in question to ensure high-, or low-involvement. Below the message, either the deceptive, or non-deceptive graphs were shown. Below the graph, the forced choice of reopening, or keeping the economy closed is presented.

After viewing the graph (one participant only views one graph in the process of the study) the participants answer questions regarding the graph to assess understanding of the graph shown. After the questions regarding the graph understanding and their self-assessment, the participants will answer the questions from the Graph Literacy Scale. This is the last step to avoid any learning effect that would take place during this stage and would influence the outcomes of the understanding of the graphs shown in the conditions.

The time required for the questionnaire varied greatly across the participants. The minimum time required for the questionnaire was 231 seconds while the participant who took longest to reply to the questionnaire took 212,594 seconds to respond to it. This leads to an average duration of 3,558.82 seconds (SE = 22,070.62). The median duration of 455 seconds, while the mode of the duration was 400 seconds.

Statistical treatment

A two-way ANOVA was performed to assess the significance of the effects of involvement (high/low) and for the effects of deception (deceptive/non-deceptive). The two-way ANOVA was also used to search for possible interaction effects between involvement and deception. The mean scores of the percentage of correctly answered understanding questions in different conditions were computed and compared to one another.

To assess if there is an alternative explanation for possible results, the graphicacy score has been analyzed. First the correct responses to the graphicacy questions were combined to a percentage of correct responses. Next, a Pearson correlation was computed to assess for possible correlations between the percentage of correctly answered graphicacy questions and the percentage of correctly answered understanding questions.

To compare if the manipulation of the involvement condition was successful the scores of the manipulation checks were combined. An independent samples t-test was performed to compare the means of the two involvement-conditions.

Results

A two-way analysis of variance with country of graph (involvement, high/low) and graph type (normal/deceptive) to assess possible effects and interaction effects on understanding.

When testing for the effect of the country of the graph on understanding, the high-involvement group had a higher percentage of correct understanding (M = 0.78, SE = 0.21%) than the low-involvement group (M = 0.78, SE = 0.21). The effect was non-significant (F (1, 166) < 1).

When testing the effect of graph type on understanding, the understanding for the deceptive graph was higher (M = 0.80, SE = 0.19) than for the normal graph (M = 0.77, SE = 0.22). The effect of graph type on understanding was not significant either (F (1, 166) = 1.18, p = .279).

Within the high-involvement condition of country of graph, the deceptive graph was understood better (M = 0.81, SE = 0.19) than the normal graph (M = 0.77, SE = 0.23). Within the low-involvement condition of country of graph, the deceptive graph was also understood better (M = 0.80, SE = 0.20) than the normal graph (M = 0.77, SE = 0.22). No significant interaction effect of country of graph and graph type has been found (*F* (1, 166) < 1).

To evaluate other possible findings, correlation analysis has been performed between graphicacy understanding, however no significant correlation between graphicacy and understanding has been found (*r* (170) = .13, *p* = .104).

An independent samples t-test showed a significant difference between the high-involvement and the low-involvement conditions regarding the perceived relevance of the information displayed to the viewers (*t* (168) = 10.47, *p* < .001) . Participants who were shown the graph depicting the Netherlands (high-involvement) were shown to be more involved (*M* = 4.45, *SD* = 1.40) than the participants who were shown the information about Zambia (low-involvement) (*M* = 2.41, *SD* = 1.14).

Conclusion

The results of the manipulation of involvement on graph understanding were insignificant. The results of the manipulation of deception on graph understanding were also insignificant, same as the interaction effect between the two manipulated factors. Thus, it was not possible to find answers to the research questions in the current study. The hypotheses of the effects of involvement or deception in visualizations on graph understanding can therefore neither be confirmed, nor rejected due to the insignificant results.

The results of a correlation between graphicacy, and of correctly answered understanding questions did not yield any significant results. The graphicacy score of participants is thus unlikely to have had any significant influence on the results.

Finally, the analysis of manipulation checks led to the significant result that participants in the high-involvement condition were more involved than in the low-involvement condition. Because the involvement condition was manipulated using indications of location, this is a confirmation of the findings of Peck et al. (2019), who found that viewers are more involved when data visualizations show an indication of location, they have a personal link to. This was the case here as the personal link was created through displaying “The Netherlands” to people residing in the Netherlands, while lowering involvement through displaying “Zambia” above the graph and in the accompanying texts.

Besides confirming the effects of involvement, and the possible success in manipulating involvement, no other significant results have been found. This leads to the conclusion that future research in the same domain requires an improved or different research approach to further investigate the dynamics between involvement, and deceptiveness of visualizations.

Discussion

In this study, visual communication tools in the form of manipulated bar stock graphs were used to search for possible influences of deception and of involvement on graph understanding of viewers. The factor of deception was manipulated by distorting the axes of the graph in one condition, while keeping the axes in the other condition non-distorted. The factor of involvement was manipulated by portraying either the Netherlands or Zambia as location in which the data was gathered, accompanied by further messages which point towards the Netherlands or Zambia as focus point of the research.

The possible confound of graphicacy (that is the ability to read, understand, and interpret graphs; Okan et al., 2016) on the percentage of questions answered correctly was not found to be present.

Finally, through manipulation checks, the findings of Peck et al. (2019) were confirmed. The findings showed that indication of location does lead to increased involvement if personal links exist.

One likely reason for the insignificant results of the manipulations of involvement and deception on understanding is likely the “conceptual question” as brought up by Peck et al. (2019). According to Johnson and Eagly (1989), increased involvement would to deeper processing of the information. In this study, involvement was manipulated by adapting the titles of the graphs, the description of the (fictional) situations, and the location indicating parts of the questions. It seems that the influence of the conceptual question asked in the questionnaire was stronger than the influence of the manipulation of the condition. It is likely that viewers intentionally directed more attention towards the graphs due to the perceived importance of those that required reading off the charts and calculating differences using the graphs as point of reference.

That the participants were able to read off the graphs they viewed at any point in time could also possibly be a factor in the non-significant effects of the deception-conditions: In the study of Roser (1990), the participants were exposed to information in one time frame. The involvement-manipulation in the study by Roser (1990) led to either increased recall, or in case of the less involved group, to decreased recall, depending on if they processed the information deeply, or shallowly. When asking questions concerning the data, while allowing the participants to look at the data in question, the results will be more accurate, also due to the influence of the conceptual question as introduced by Peck et al. (2019).

The reasons for the deception condition not leading to significant results on understanding might be a too soft manipulation of the factor. In the study of Sun et al. (2012), the look of the graph was changed completely, while the data portrayed remained the same. In the current study, the changes performed on the graphs was less severe – while the rise of the graph in the deceptive condition was visually steeper than in the normal graph, the overall image was not changed significantly.

In conclusion, the manipulation of the factor deception was not strong enough to show an effect. In the future, the deceptiveness of the graphs could be improved. The goal of the study was to investigate visual presentation and deceptiveness of visual presentations of data; However, the labels which were intended to help and guide participants through the tasks might have simplified the tasks too much. This could have changed the tasks to be more of a graph reading task. In the case of the factor involvement, the manipulation was weaker than the influence of the conceptual questions which might have negated the effect of the manipulation on attention.

Future research might pick up where this study reached its limitations: The influence of the conceptual question compared to the influence of other factors that change the involvement of participants might be an interesting topic. Such research could also lead to more insights into how the process of understanding visual communication works, and which factors play which role.

An improved design of the current study might include a difference in time between receiving the information and responding to questions. This would test processing in the form of recalling the information received. Such a design would assess possible influences of deep or shallow processing through measuring accuracy in recall, similar to the study performed by Roser (1990). This eliminates the influence of the conceptual question on attention directed towards the information, as the questions would be unknown before the information is received.

Overall, the current study offered some insights into the greater theory of visual communication understanding, even though the insights were not the effects anticipated. It offers new perspectives on the degree to which attention towards visual communication is influenced by the conceptual message in comparison with general involvement, even though further research is required. In that sense, the study should be improved and continued to further the insight into the greater topic of visual communication.

References:

Boers, M. (2018). Designing effective graphs to get your message across. *Annals of the rheumatic diseases*, *77*(6), 833-839.

Brown, S. M., Culver, J. O., Osann, K. E., MacDonald, D. J., Sand, S., Thornton, A. A., ... & Weitzel, J. N. (2011). Health literacy, numeracy, and interpretation of graphical breast cancer risk estimates. *Patient education and counseling*, *83*(1), 92-98.

Durand, M. A., Yen, R. W., O'Malley, J., Elwyn, G., & Mancini, J. (2020). Graph literacy matters: Examining the association between graph literacy, health literacy, and numeracy in a Medicaid eligible population. *PloS one*, *15*(11), e0241844.

Frewer, L. J., Howard, C., Hedderley, D., & Shepherd, R. (1997). The elaboration likelihood model and communication about food risks. *Risk analysis*, *17*(6), 759-770.

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy: A cross-cultural comparison. *Medical Decision Making*, *31*(3), 444-457.

Garcia-Retamero, R., & Galesic, M. (2010). Who proficts from visual aids: Overcoming challenges in people's understanding of risks. Social science & medicine, 70(7), 1019-1025.

Johnson, B. T., & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. Psychological bulletin, 106(2), 290.

Liberman, A., & Eagly, A. H. (1989). Heuristic and systematic information processing within and beyond the persuasion context. Unintended thought, 212-252.

Maichle, U. (1994). Cognitive processes in understanding line graphs. In Advances in psychology (Vol. 108, pp. 207-226). North-Holland.

Nayak, J. G., Hartzler, A. L., Macleod, L. C., Izard, J. P., Dalkin, B. M., & Gore, J. L. (2016). Relevance of graph literacy in the development of patient-centered communication tools. Patient Education and Counseling, 99(3), 448-454.

Nelson, W., Reyna, V. F., Fagerlin, A., Lipkus, I., & Peters, E. (2008). Clinical implications of numeracy: theory and practice. Annals of behavioral medicine, 35(3), 261-274.

Okan, Y., Galesic, M., & Garcia-Retamero, R. (2016). How people with low and high graph literacy process health graphs: Evidence from eye-tracking. Journal of Behavioral Decision Making, 29(2-3), 271-294.

Okan, Y., Stone, E. R., & Bruine de Bruin, W. (2018). Designing graphs that promote both risk understanding and behavior change. Risk Analysis, 38(5), 929-946.

Padilla, L. M., Creem-Regehr, S. H., Hegarty, M., & Stefanucci, J. K. (2018). Decision making with visualizations: a cognitive framework across disciplines. *Cognitive research: principles and implications*, *3*(1), 1-25.

Pandey, A. V., Rall, K., Satterthwaite, M. L., Nov, O., & Bertini, E. (2015, April). How deceptive are deceptive visualizations? An empirical analysis of common distortion techniques. In Proceedings of the 33rd annual acm conference on human factors in computing systems (pp. 1469-1478).

Peck, E. M., Ayuso, S. E., & El-Etr, O. (2019, May). Data is personal: Attitudes and perceptions of data visualization in rural pennsylvania. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-12).

Peebles, D., & Ali, N. (2009). Differences in comprehensibility between three-variable bar and line graphs. In Proceedings of the thirty-first annual conference of the cognitive science society (pp. 2938-2943). Lawrence Erlbaum Mahwah, NJ.

Petty, R. E., & Cacioppo, J. T. (1984). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of personality and social psychology*, *46*(1), 69.

Planinic, M., Ivanjek, L., Susac, A., & Milin-Sipus, Z. (2013). Comparison of university students' understanding of graphs in different contexts. Physical review special topics-Physics education research, 9(2), 020103.

Romano, A., Sotis, C., Dominioni, G., & Guidi, S. (2020). Covid-19 data: The logarithmic scale misinforms the public and affects policy preferences.

Roser, C. (1990). Involvement, attention, and perceptions of message relevance in the response to persuasive appeals. Communication Research, 17(5), 571-600.

Spiller, S. A., Reinholtz, N., & Maglio, S. J. (2020). Judgments Based on Stocks and Flows: Different Presentations of the Same Data Can Lead to Opposing Inferences. Management Science, 66(5), 2213-2231.

Sun, Y., Li, S., Bonini, N., & Su, Y. (2012). Graph-framing effects in decision making. Journal of Behavioral Decision Making, 25(5), 491-501.

Stone, E. R., Reeder, E. C., Parillo, J., Long, C., & Walb, L. (2018). Salience versus proportional reasoning: Rethinking the mechanism behind graphical display effects. Journal of Behavioral Decision Making, 31(4), 473-486.

Appendix B – Complete questionnaire and material used in the Bachelor Thesis including stimuli and translation into English as well as Dutch (see next page

Appendix B – Statement of own work

Student name:
Student number:

PLAGIARISM is the presentation by a student of an assignment or piece of work which has in fact been copied in whole or in part from another student's work, or from any other source (e.g. published books or periodicals or material from Internet sites), without due acknowledgement in the text.

DECLARATION:
a. I hereby declare that I am familiar with the faculty manual (https://www.ru.nl/facultyofarts/stip/rules-guidelines/rules/fraud-plagiarism/) and with Article 16 "Fraud and plagiarism" in the Education and Examination Regulations for the Bachelor's programme of Communication and Information Studies.
b. I also declare that I have only submitted text written in my own words
c. I certify that this thesis is my own work and that I have acknowledged all material and sources used in its preparation, whether they be books, articles, reports, lecture notes, and any other kind of document, electronic or personal communication.

Signature:

Place and date: