Do you care? The influence of topic relevance on the understanding of time series data presented as a stock and flow

Radboud Universiteit

Radboud University Nijmegen, the Netherlands

Study degree: Bachelor International Business Communication

Supervisor: Peter de Swart

Student: Florian Verschuren

Student number: S1019580

Date and place of submission: 06-06-2021

Abstract

The way data is presented influences the understanding of this data. Time-series data could be presented as a stock (absolute values in a period of time) and as a flow (difference in absolute values over a period of time). By conducting an online experiment in the form of a questionnaire, it has been found that time series data presented as stock graphs are understood better than time-series data in the form of flow graphs but this might be due to certain graphical features and the communicational goal that is to be fulfilled with the graph. Moreover, the present article investigated the role topic relevance could play on the understanding of stock and flow data. The results showed that the relevance of the topic is depended on several factors and that it does not significantly change the understanding. Thus, the type of data visualisation is found to influence the understanding of the graph whereas the relevance of a topic appears to have no significant effect. Finally, limitations and recommendations for future research are discussed.

Introduction

Data is a keyword in the current information age. Data is being shared by ever developing forms of technology and helps us to make sense of the world. In every occupation some sort of data is being distributed in order to be able to function properly, think of news anchors, financial managers and doctors. Without data it would be very much impossible to perform any of these occupations. One way of sharing data is by simply telling or writing it towards the addressee. However, studies have shown that merely textual data is not always understood correctly (Lipkus, Samsa, & Rimer, 2001; Peters et al., 2006; Reyna & Brainerd, 2008). Apart from the practical inconveniences of misunderstanding information, having an accurate understanding of data in situations that are related with health or risk communication could even be vital. Therefore, a better option for presenting data needs to be found and according to a broad area of research is data visualisation a suitable alternative for textual data in order to get an accurate understanding of the information.

One way of visualizing data is by creating graphs. Well-designed graphs have the potential to be a memorable, vivid and easy-to-understand way of depicting quantitative data (Larkin & Simon, 1987; Shah, Freedman & Vekiri, 2005). Therefore, graphs are being used extensively in popular printed media, textbooks and scientific journals (Zacks et al., 2002; Shah et al., 2005). Moreover, their prevalence has more than doubled in newspapers between 1984 and 1994 according to Zacks et al. (2002).

A great deal of research has provided results indicating that information presented visually, in the form of graphs, is better understood than textual data (Garcia-Retamero & Galesic, 2010; Zikmund-Fisher et al., 2008; Petrova, Garcia-Retamero & Cokely, 2015; MacDonald-Ross, 1977). In Garcia-Retamero and Galesic (2010) an experiment was conducted in which the degree of risk understanding of numerical data was compared between visual data, in the form of an icon array and a bar graph, and plain textual data. The results clearly showed that people had a better understanding of certain fictional risks when presented with visual data than was the case when they only saw numerical data written as a text. Moreover, in Petrova, Garcia-Retamaero and Cokely (2015) two experiments were conducted on the effect of graphical information on the comprehension of the risks of cancer screenings as opposed to the same data presented visually. The results indicated that the participants understood the risks better when presented with the visual data. However, the results also indicated that previous beliefs about the topic may interfere with the understanding of the data.

The processes involved in understanding the data visualized in a graph could be explained by the model of graph comprehension presented in Pinker (1990). According to Pinker's model, one first makes a visual description of the graphical elements (e.g. colours, types of lines, the symbols) by making use of several constraints such as the Gestalt principles of similarity, good continuity and proximity. More specific, the Gestalt laws of relevance here are the laws of prior experience and grouping (Pinker, 1990). The visual description, then, is being used to create conceptual messages which are propositions about the graph variables. Lastly, one aims to draw conclusions from these conceptual messages, making use of high-level inferential processes.

Pinker (1990) argues that one type of graph could be understood better than another type depending on the data that is to be extracted from the graph. This could partly explain the differences in comprehension between different types of graphs found in prior research. In Zacks and Tversky (1999) an experiment was conducted in which the participants had to describe what is shown in the graph that was presented to them. The participants saw either a line or a bar graph but both depicted the same data. The results clearly showed that people understood the data in terms of discrete data points when presented as a bar graph and as a trend when presented as a line graph, indicating the graph format has an influence on understanding.

Besides graph format, the way in which data is presented could also be a factor of relevance to understanding. Especially in the field of risk communication this phenomenon is observed

(Stone et al., 2003; Hawley et al., 2008; Okan et al., 2015; Galesic, Garcia-Retamero & Gigerenzer, 2009). In research by Stone et al. (2003) participants were presented with either a graphical display that depicted the total number of people at risk of periodontal disease when using a certain type of toothpaste along with the people who already were affected by the disease due to the toothpaste, and a graph that merely depicted the amount of people that were affected due to the product. The results showed that the participants who were presented with a graph depicting the people who are at risk and the people that were already harmed, had a better understanding of the risk than those who merely saw the affected people.
However, evidence outside the field of risk communication is also available. In Carswell, Emery and Lonon (1993), experiments were conducted in which the participants had to provide written interpretations of single-function line graphs which they got to see spontaneously, that is, without any indication of what to pay attention to. Certain graphical elements were manipulated in order to change the structural complexity of the visualization. The results showed that multiple graphical features could determine the understanding of the graph but that the largest effect was found for the number of trend reversals. Other graphical elements that played a role were, the amount of data points, symmetry of the graph and the departures from linearity.

Most studies discussed so far are focussed on types of data depicting merely one moment in time. However, data about multiple consecutive periods in time (e.g. time series data) is used in 75% of the graphics presented in magazines and newspapers (Tufte, 2001) and is therefore rather common. When time series data is to be depicted this is often done in the form of a stock or flow. The first one depicts the extent of the target quantity at different periods and data depicted as a flow shows the change of the target quantity between those periods (Spiller, Reinholtz, & Maglio, 2020). Prior research has provided findings indicating that there might be a difference in data understanding when presented as a stock or as a flow. In Cronin et al. (2009) and Booth Sweeney and Sterman (2000), an experiment was conducted in order to test whether people are able to translate between stock and flow data. During the experiment, a data visualisation was presented in which stock information was given about the amount of people entering and leaving a fictional store in a time period of thirty minutes with on the y-axis the amount of people and on the x-axis the duration in minutes. In order to interpret the graph the participants had to translate the stock information into flow data via the process of accumulation. The participants were asked questions in which they had to indicate, among others, after how many minutes most people were in the store and when the least. In

order to do so, the participants had to accumulate the stock data. The results of both studies showed that people often gave incorrect answers, which indicates that the participants have difficulty in understanding the process of accumulation and therefore have difficulties in translating between stock and flow data.

Based on these findings, Spiller et al. (2020) conducted a total of six experiments to test whether people make different judgements about the job market situation before and after Obama's presidency if the same data is visualized as a stock or a flow (figure 1). In the first experiment the participants saw either flow or stock data. The flow data showed month-to-month changes in the amount of private-sector jobs in the United States in a five year period, whereas in the stock condition, the participants saw the total amount of private-sector jobs per month. The periods in which Obama and his predecessor Bush were president were coloured differently. In order to verify whether the participants understood the graph correctly, they were asked when the amount of private-sector jobs started to rise and when the rate of private-sector job loss started to slow based on the chart. With respect to the first question, it appeared that overall the participants, in both conditions, were correct in describing when the amount of jobs started to rise. Although, there was a small difference, as slightly more people were correct when assigned to the stock condition than to the flow. When it comes to the second question, there was more variation between the conditions. In the flow condition, 76% of the participants indicated that the slowing of the job loss rate started between January 2009 and September 2009 (which was right) whereas merely 40% did so in the stock condition. Moreover, another 40% of the stock condition gave "somewhere in 2008" as an answer. These results might indicate that people have a more accurate understanding of flow data than stock data. Thus, as described above, the type of data could also have an influence on understanding when it comes to time series data.

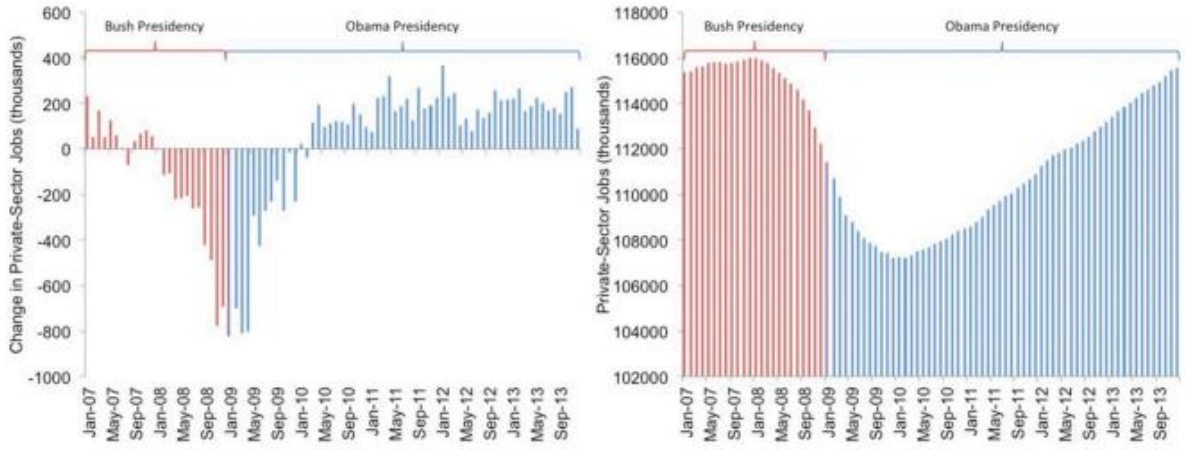


Figure 1: The materials used in experiment 1 in Spiller et al. (2020). With the flow condition left and the stock condition right.

Moreover, besides graphical format and type of data, it is possible that engagement in reading the graph also plays a role in graph understanding. When one is highly engaged to read a graph, this may lead to this person putting more effort (e.g. dedicating more processing capacity) in understanding it and therefore the chances of understanding the data correctly might be higher than when one is not engaged (Petty, Wegener & Fabrigar, 1997; Petty & Wegener, 1998). This can be explained using the Elaboration Likelihood Model (ELM) by Petty and Cacioppo (1980). According to the ELM, one is able to process stimuli via two mental paths or "routes". The first route is the so-called "central route", which involves information processing in a rather exhaustive manner. One makes use of this route when they are motivated to a great extent to understand the information presented (e.g. high engagement). When one is for example very much interested in politics, this person is likely to read the election programmes of multiple parties, leading to the arguments being processed via the central route, resulting in this person carefully weighing the arguments. If one is to a lesser extent interested in politics but is still presented with the election programmes, then this person is more likely to quickly scan through the programmes, looking for heuristic cues that could determine one's opinion (such as the names of well-known politicians or political parties alongside the arguments), rather than carefully reading all the texts. The latter is an example of one taking the heuristic processing route, in which a person is not very much motivated to process the information presented and is likely to solely scan for relevant names or numbers. In Cronin et al. (2009), it was hypothesized that motivation to give the correct answer could influence the understanding of stock and flow graphs. In their experiment, motivation was manipulated by letting the participants hand in their answer form to the researcher who would then correct the sheet. The participants were asked to do the task again if they had given one or more incorrect answers, until they had answered all questions

correctly. These results were compared to the no motivation condition. The results did, however, not indicate a significant difference in performance. Nevertheless, it might have been that motivation (e.g. engagement) was not manipulated sufficiently.

One of the most important factors in engagement might be the relevance of a topic to the receptor. Results in favour of this notion can be found in a variety of studies, such as the one of Assor, Kaplan and Roth (2002). In this study, the effects of different teacher behaviour styles were tested on children. The results showed that when teachers actively explained their students why certain topics were of relevance to the children this led to more engagement in the subject itself compared to teacher behaviours in which subjects and tasks were forced upon the children. These findings indicate a positive relationship between (perceived) topic relevance and engagement.

Similar results with respect to this proposed relationship could also be found in the field of visual communication. Peck, Ayuso and El-Etr (2019) conducted an experiment in which people in rural Pennsylvania were asked, among others, to choose out of 10 different graphs depicting drug use in the United States which one was the most useful and also why they thought so. The results show that people with prior experience to a certain topic, gravitated towards the graphs that depicted this topic. Moreover, the participants often told their personal stories related to the topic of the graph, whereas graphs that were about less relevant themes were not or barely discussed. Furthermore, they also found that with respect to data visualizations containing maps, the participants thought the map was more useful if it included their current place of living or their former, as they immediately searched for this place on the map. Thus, it might as well be the case that people reach higher levels of engagement when data is provided about places that are meaningful to them, as it appears to be that data becomes more relevant if it is about one's current or former place of living.

Data in favour of the proposed relation between topic relevance and graph engagement can also be found in the field of the social sciences. In Kennedy and Hill (2018), multiple focus groups were given diaries in which they had to indicate, on an already provided template, if, when, where and how they encountered any visualizations in the week before their focus group meetings. They also had to describe what their first impressions were, how they felt at the time and what they thought about the visualisation. Then, in the focus group meetings, the participants were shown several data visualisations and were asked to indicate for each graph whether they liked and learned anything from them. Their results clearly showed that people were more engaged in reading the graph when the topic was of relevance to them than when this was not the case. To illustrate, there were several graphs that had a migration related topic

and these graphs were discussed mostly by the participants that had prior experience with migration, whereas the people that had no direct relation to this topic were to a lesser extent or not at all engaged with these graphs. They conclude by arguing that topic relevance is a relevant factor when it comes to engagement with visualisations.

When taking a look at the study by Spiller et al. (2020), one could argue about the relevance of the topic (the situation of the job market in the United States in recent years). Although, this might seem a universal interesting topic (for Americans), it could be the case that most participants had a very stable job at that moment and therefore had low interest in the creation or loss of jobs. This might have had an effect on the engagement to read the graph and therefore on the understanding of the depicted data. The effect would be that the participants were not very much engaged and therefore did not carefully read the stock and flow graphs which had led to the low results in understanding. It might therefore be the case that when the topic actually is of relevance to the participants, this leads to more engagement with the graph and therefore to a better understanding of the graph

The present study

The present research focuses on the influence of topic relevance on the understanding of visual stock and flow data. More precise, the participants of the experiment will be presented time series data about the COVID-19 infections in the Netherlands (e.g. high relevance condition) and the African country Zambia (e.g. low relevance condition) in either a stock or a flow depiction.

In Spiller et al. (2020), the results indicated that there were differences in the understanding of flow and stock data. The greatest difference occurred at the first question: "Based on the chart above, when did the rate at which the United States was losing private-sector jobs start to slow?" and indicated that people were less accurate at understanding data when presented as a stock than when presented as a flow. Based on this finding the following hypothesis is formulated:

H1: People will understand time series data presented as a flow better than time series data presented as a stock.

With respect to the influence of topic relevance, it is predicted that people will be more engaged in reading both types of graphs when the topic of the graph is of relevance to them. This, because it is expected that when people find a topic to be relevant, they will

process this information via the central cognitive route according to the ELM. This central route processing will lead to a more careful consideration of the data presented. Based on the findings by Peck, Ayuso and El-Etr (2019), it is expected that people find data about a country where they currently live in or live close to, more relevant than the same data about a country they have no bond with. Therefore, it is expected that people find COVID-19 infection data about the Netherlands more relevant than the infection data of Zambia and are therefore more engaged in reading the graph about the Netherlands. As this is the case, it is expected that people will have a better understanding of the data with a graph that is of high relevance to them, because this might lead to central route processing, than when this is not the case. Specifically, this leads to the following hypothesis:

H2: People will have a better understanding of time series data that is of high relevance than time series data that is of low relevance.

Furthermore, because it is expected that people will be more engaged to read a graph that is of relevance to them, it is predicted that topic relevance has a positive influence on the understanding of data being presented in either a flow or a stock format. This leads to the third hypothesis:

H3: *The difference in understanding between stock and flow time series data will decrease if the topic relevance is high.*

Method

Materials

The independent variables in the present article were topic relevance, which could either be high or low and type of visualization which could be either in the form of a stock or flow. (Figures 2,3 and Appendix B)

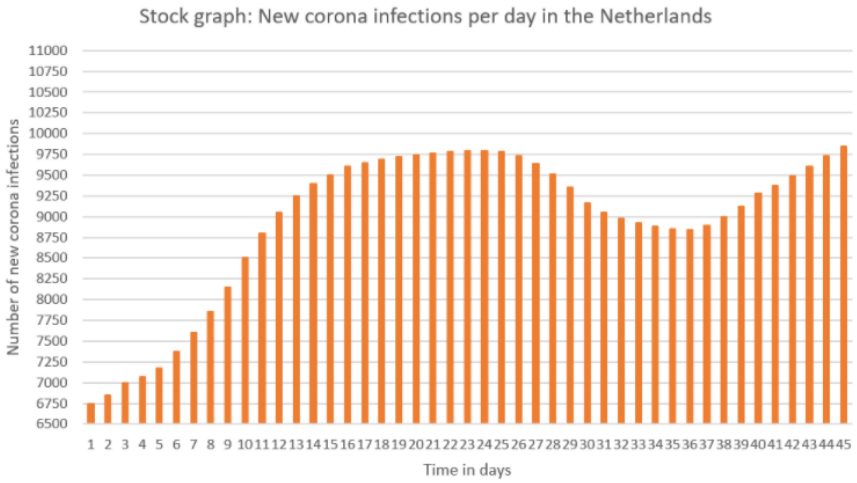


Figure 2. One of the stock graphs used in the experiment.

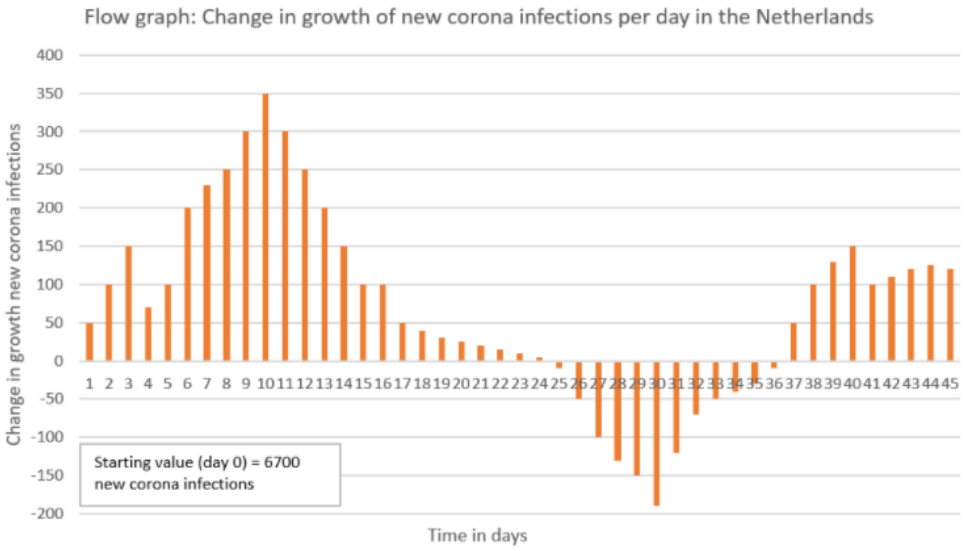


Figure 3. One of the flow graphs used in the experiment.

All graphs were about new COVID-19 infections per day in a particular country. For the low relevance conditions infection data of the African country Zambia was picked because this country is unlikely to be of relevance to most people living in or close to the Netherlands, as this country is barely mentioned in Dutch news and because the two countries do not share any colonial history with each other. With respect to the high relevance conditions, COVID-19 data of the Netherlands was picked as this data is of direct relevance to Dutch people. The names of the countries were mentioned above the graph in order to make sure that the participants keep the country in mind while reading the data. The data used was not the actual

COVID-19 infections data of the two countries, as the same pattern for both countries was incorporated in order to facilitate comparisons. Moreover, in order to make more valid comparisons, the data had a somewhat similar pattern as the graphs used in experiment 1 of Spiller et al. (2020), that is, an increasing trend followed by a decrease and another increase (figures 1,2,3 and Appendix B). It was made sure that the title gave away what type of graph the participants were looking at and also what the data actually represented, as this reduced the chance of participants answering without knowing how to interpret the graph. There were no further differences between the graphs except that, for the sake of clarity and readability, the intervals on the y-axis for the stock condition were 250 and 50 for the flow condition and that the y-axis ranged from 6.500 to 11.000 and from -200 to 400 respectively. Moreover, in the flow condition, there was a box indicating the number of infections on day 0, as this information is needed in order to answer the questions. Furthermore, as most of our participants were expected to be Dutch and in order to be sure that everyone would understand the questions being asked, there were two versions of the questionnaire: Dutch and English. After the experiment, a manipulation check was conducted in order to test how relevant people perceived the data to be.

Participants

In total there were 262 respondents who filled in the questionnaire. The data was collected by multiple researchers who were all investigating their own specific research questions which had led to a more extensive questionnaire and also to two more experimental conditions. As these conditions were not part of the present study, the participants in those conditions were removed from the data file used in the current research. Other requirements were that the participants had to live in the Netherlands, were older than eighteen and had filled in the entire survey. Moreover, for the manipulation check questions, in two of the analysed conditions the questions were accidently not marked as mandatory, leading to one participant not filling in those questions. This person was therefore removed. After removing the participants that did not met one or more of the requirements mentioned above, a total of 162 participants were left of whom the responses were actually used in the analysis.

Within those 162 participants, there were 60 (37.0%) males, 98 (60.5%) females, three (1.9%) people who identified themselves differently and one (0.6%) participant who did rather not share their gender. The ages of the participants ranged from 18 to 66 (M = 27.85, SD = 12.73). With respect to the educational level of the participants, most people's current or

last completed education was of a bachelor university level and the levels ranged from high school to PhD. A total of 24 participants filled in the questionnaire in English whereas 138 participants took the Dutch survey.

Design

The experiment had a 2x2 between subjects design. Thus, there were four different groups all receiving different stimuli. That is, all participants were presented with a graph depicting COVID-19 infection rate data in a certain country. This graph could either visualize the data as a stock or flow and could be about Zambia (e.g. low relevance condition) or about the Netherlands (e.g. high relevance condition).

Instruments

The variable understanding was operationalized by letting the participants fill in an online questionnaire which was created in Qualtrics. As mentioned before, multiple researchers with all their own area of interest, have constructed the questionnaire. Therefore, there were more questions than that have actually been used in the analysis. Merely, the questions that have been used in the analysis are discussed. The questionnaire contained five questions measuring understanding and all these questions were of relevance to the present study. As most of our participants were expected to be Dutch and in order to be sure that everyone understands the questions being asked, there were two versions of the questionnaire: Dutch and English.

Four of the understanding questions were created by the researchers and were expected to give an insightful indication of whether the participants understood the graphs. In chronological order these questions were: (1) “How many new covid cases were registered on day 3?”, (2) “Look at day 3 and day 4. Which day shows the biggest increase in new infections compared to the day before?”, (3)“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?” and (4) “What is the difference in the number of new infections between day 12 and day 13?”. The fifth understanding question is derived from Spiller et al. (2020) but adapted to the present experiment: (5) “On which day did the decrease of covid-19 infections start to slow down?”.

Questions one, four and five were open-ended questions where the participants could type any number in the answer box. With respect to the second and third question, there were

two answer options which were either “day 3” or “day 4” or either “the period from day 1 to day 10” or “day 11 to day 20” respectively.

The manipulation check consisted of three adapted items derived from the study by Frewer et al. (1997) and indicated how relevant the people the graph, the infection rate and the COVID-19 regulations of the given country perceived to be. The items consisted of the following claims: “The graph I saw is very relevant to me personally”, “The positive COVID-19 tests in [Zambia/the Netherlands] are very relevant to me personally” and “The COVID-19 regulations in [Zambia/the Netherlands] are very relevant to me personally”. The participants were asked to indicate to what extend they agreed to each claim on a 7-point Likert scale that ranged from: 1 (strongly disagree) to 7 (strongly agree). The reliability of “Manipulation check” comprising three items was acceptable: $\alpha = .78$. Consequently, the mean of all three items was used to calculate the compound variable “total manipulation check”, which was used in further analysis.

Procedure

The participants were recruited by sending invitations to participate in the experiment via Qualtrics making use of various social media channels (e.g. Facebook, Instagram and WhatsApp). As the experiment in the present research is part of multiple studies that will be using the same data, multiple researchers have distributed the questionnaire via the indicated channels.

There were no rewards or anything similar promised to the participants. Neither were there any restrictions on whether the survey should be made alone or not. Furthermore, there was no additional information about the aim of the study provided as this could have had an impact on the involvement of the participants with respect to the topic and the questions.

When one received the invitation, the receiver was be able to click on it, directing them to the questionnaire. First the participants saw a text in which they were thanked for their participation. The next page contained information about the content of the questions, where one could find their progress on the screen, who conducted the research, ethical principles were explained (such as that the data is anonymous and that they will only be used for this particular study) and the participants were asked to give a confirmation of agreement for the informed consent. When going to the next page, the participants were asked about their

age, gender, country of living and their last completed or current level of education. Next, after answering some questions related with a variable unrelated to this study, the participants were asked the five questions measuring understanding. The graph was depicted on top of the screen with the questions and answer boxes located underneath. The participants were also able to zoom in and out on the graph and to scroll up and down the page to take a look at the graph or read the questions again. Once finished with the understanding questions, the participants were presented with three questions that were part of a manipulation check. These questions measured how relevant the topic of the previous questions were to the participants. When the participants finished the questionnaire they were directed to a final page where they were once again thanked for their participation. For all participants the same procedure was be used. It took around 10 minutes to complete the questionnaire.

Statistical treatment

In order to test for a main effect of type of visualization and topic relevance on understanding a two-way ANOVA was conducted, this analysis is also used to test the interaction of the independent variables. The variable understanding consists of the amount understanding questions that were answered correctly, ranging from 0 to 5. Furthermore, an independent samples t-test was conducted between topic relevance and total manipulation check in order to test whether the variable topic relevance was manipulated sufficiently.

Results

A two-way analysis of variance with types of visualisation and topic relevance as factors showed a significant main effect of types of visualisation on understanding ($F(1, 158) = 117.40$, $p < .001$). There was a significant main effect of types of visualisation on understanding, which in this case meant that it was shown that people understand a time series graph better when presented as a stock ($M = 4.15$, $SD = .99$) than when presented as a flow ($M = 2.53$, $SD = .90$). Topic relevance was not found to have a significant main effect on understanding ($F(1, 158) < 1$, $p = .611$). The interaction effect between type of visualisation and topic relevance was not statistically significant ($F(1, 158) < 1$). (See table 2 for the amount of participants, the means and the standard deviations per condition and in total).

Table 2. Number of participants (n) Means (*M)* and Standard deviations (*SD)* for the understanding per condition and in total. The means are the averages of the amount of understanding questions that were answered correctly, ranging from 0 to 5.

	Involvement				Total	
	Low		High			
Type of visualization	n	*M*(*SD*)	n	*M*(*SD*)	n	*M*(*SD*)
Flow	36	*2.44*(*.81*)	40	*2.60*(*.98*)	86	*2.53*(*.90*)
Stock	46	*4.15*(*.97*)	40	*4.15*(*.98*)	86	*4.15*(*.99*)
Total	82	*3.40*(*1.24*)	80	*3.38*(*1.27*)		

A manipulation check for topic relevance was conducted in order to verify whether the participants indeed perceived the low relevance condition to be of less relevance than the high relevance condition. An independent samples t-test showed a significant difference between people in the low relevance condition and those in the high relevance condition with regard to the compound variable total manipulation check ($t(160) = 10.08$, $p < .001$). Participants in the low relevance condition ($M = 2.27$, $SD = 1.13$) were shown to perceive the overall topic to be of less relevance to them than participants in the high relevance condition ($M = 4.24$, $SD = 1.35$).

Discussion

The goal of the present research was to investigate whether one's relevance to the topic of a graph has an influence on the understanding of the same graph. Moreover, it was aimed to test whether this possible influence variated between two different types of visualizing time series data (e.g. as a flow or a stock) and if there is a difference in understanding between those two types of data visualisation.

The present study did not find support for the notion that time series data presented as a flow are understood better than the same data presented as a stock. Moreover, the current research did not find results supporting the hypothesis that people will better understand time series data that is of relevance to them than time series data that is irrelevant to them. Finally, the present research did not find evidence in favour of the notion that the relation between

type of visualisation and understanding was moderated by the relevance of the topic of the graph.

As mentioned above, the results do not indicate a better understanding of time series data presented as a flow than time series data in the form of a stock. In fact, the results even indicated that the participants understood the stock graphs better. This finding is also in contradiction with the finding in experiment 1 of Spiller et al. (2020), of that the flow graph was understood better than the stock graph. A possible explanation for this difference might be that the questionnaire in the present research allowed the participants to zoom in on the graph, whereas this option is not mentioned in Spiller et al. (2020); this might have made the graphs of the present study easier to read. Moreover, it is likely that the stock graph used in the present study was easier to read than the one used in Spiller et al. (2020). This might have been the case as the current stock graph had fewer data points, showed much lower values and had much smaller intervals on the y-axis. In Carswell, Emery and Lonon (1993) it is found that certain graphical features, such as the amount of data point, could influence the understanding of the graph. This indicates that a possible difference between flow graphs and stock graphs with respect to understanding could be reduced or altered by manipulating certain graphical features. More precisely, it might be that graphs that have many data points and large intervals on the y-axis are better understood depicted as a flow than when visualized as a stock and that this effect reduces and possibly reverses due to manipulations of such graphical features.

Moreover, the possible difference between the understanding of stock and flow graphs might also vary depending on what one exactly wants to show with the graph. If one for example has the goal to inform people about the total amount of COVID-19 infections in a country, it is better with regard to understanding the graph, to present the data as a stock instead of a flow. Similarly, if data about the changes of the amount of in intensive care hospitalizations is to be depicted, this could better be done in the form of a flow.
Another explanation for why different results have been found between the experiment of Spiller et al. (2020) and the present study might be that the questions used in the present experiment were different in some aspects. As Spiller et al. (2020) based their finding on one question, the present experiment consisted of five questions. The four made up questions could have been easier to answer than the question used in Spiller et al. (2020), because this question did not give any indication as to which data points the participants should look at in

order to answer the question, whereas the four made up questions did indicate this. This might have led to an overall better understanding of the stock graph.

Furthermore, the present study hypothesized that the participants would be more engaged in reading a graph when the topic is of relevance to them, which, in turn, would lead to a better understanding of the content of the graph. As already mentioned, in order to verify whether the relevance of the topic was manipulated sufficiently, a manipulation check was conducted which indicated that the participants indeed perceived the data about Zambia to be of less relevance than the data about their country of living. This is in line with the finding of Peck, Ayuso and El-Etr (2019), that people perceive data about a place that is personally important to them to be more relevant than data about other places. In their research, such a place was somewhere they are currently living or have lived. Similarly, in the present research, data about the current place of living (e.g. the Netherlands) was perceived more relevant than data about another place (e.g. Zambia). Furthermore, the manipulation check also indicated that people perceived the topic itself (e.g. COVID-19 infections rate) to be relevant. As COVID-19 is a factor that everyone is affected by, the information of the graph is also perceived to be relevant. This is in line with the proposition of Kennedy and Hill (2018) that people think that graphs depicting data about a topic they have experience with are more relevant than graphs that do not visualize such a topic.

Nonetheless, the results do also show that, although the data about the Netherlands was perceived to be more relevant, the actual indicated relevance was still not very high, in fact it was almost neutral. This might be due to the fact that the participants were informed that the data was made up. This was done in order to facilitate comparisons between the conditions but it might have had a negative influence on the relevancy and engagement and therefore also perhaps with answering the questions. This could be explained with the ELM. As the participants were informed about the fact that the data was made up, this could have led to the participants not perceiving the data to be very relevant to them and were therefore not very engaged in reading the graph. As this was the case, according to the ELM, the participants did not take a central processing route and have therefore not dedicated much processing capacity to interpret the graph. Therefore, the present study might not have been able to get people actually take the central route while reading the graph.

Moreover, it may be that topic relevance was not manipulated sufficiently as it is possible that the participants did not or barely take the country into account while reading the graph and while answering the questions, as this information was merely presented at the beginning of

the page and in the title of the graph. Therefore, the manipulation check may not have measured the actual effect of the manipulation of the variable topic relevance during the answering of the questions. Thus, as the present research did not find support for the notion that topic relevance influences the understanding of graphs, the current study does not provide support that the results in Spiller et al. (2020) would have been different if the topic of the graphs were of more relevance to the participants. Moreover, the small differences in perceived topic relevance (and therefore in engagement) between the involvement conditions and the almost neutral perceived relevancy in the high relevance condition may also explain why there has not been found a significant interaction effect between topic relevance and type of visualization.

A first limitation of this study is that the name of the country was not presented in the questions, as this may have led to the participants not taking the country in to account while reading the graph and answering the questions. Including the name of the country is therefore recommend for future research. Another limitation is that it appeared to be that the zoom function did not work for all participants while answering the questions, this may have made it harder for some participants to read the graph which, in turn, may have had an effect on the responses.

Besides the already mentioned recommendation for future research, it might also be interesting to investigate whether differences in topic relevance could influence the understanding of other common ways of data visualisation, such as pie charts and icon arrays. This, as the present research is merely focussed on time series data. It might be that topic relevance influences graph understanding depending on the type of data that is depicted. Such a study would set the path for a new area of visual communication research. Moreover, future research could be conducted to further investigate the relation between graphical features, such as the amount of data points and the length of the intervals on the y-axis, and the understanding of stock and flow graphs.

The present study has made several contributions to the existing literature on visual communication. Firstly, it had provided results indicating that time series data presented as stock graphs could be understood better than when presented as a flow. Moreover, the present study argues that this effect could differ and even reverse depending on certain graphical features. Furthermore, this research has found that people perceive data about matters they have experience with and that are about places that are meaningful to them, to be of more

relevance than data about unfamiliar topics and not meaningful places. These findings contribute new insights to the field of visual communication research and makes way for a new research direction within this field.

References

Assor, A., Kaplan, H., & Roth, G. (2002). Choice is good, but relevance is excellent: Autonomy-enhancing and suppressing teacher behaviours predicting students' engagement in schoolwork. *British Journal of Educational Psychology, 72*(2), 261–278. https://doi.org/10.1348/000709902158883

Booth Sweeney L., Sterman J.D., (2000) Bathtub dynamics: Initial results of a systems thinking inventory. System Dynam. Rev. 16(4):249–286.

Carswell, C.M., Emery, C. and Lonon, A.M. (1993). Stimulus complexity and information integration in the spontaneous interpretation of line graphs. Applied Cognitive Psychology 7:341–357.

Cronin M.A., Gonzalez C., Sterman J.D., (2009) Why don't well-educated adults understand accumulation? A challenge to researchers, educators, and citizens. Organ. Behav. Human Decision Processes 108(1):116–130

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., & Gigerenzer, G. (2009). Using icon arrays to communicate medical risks: overcoming low numeracy. *Health psychology : official journal of the Division of Health Psychology, American Psychological Association*, *28*(2), 210–216. https://doi.org/10.1037/a0014474

Garcia-Retamero, R., & Galesic, M. (2010). Who profits from visual aids: overcoming challenges in people's understanding of risks [corrected]. *Social science & medicine (1982), 70*(7), 1019–1025. https://doi.org/10.1016/j.socscimed.2009.11.031

Hawley, S., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient education and counseling, 73 3*, 448-55 .

Kennedy, H., & Hill, R. L. (2018). The Feeling of Numbers: Emotions in Everyday Engagements with Data and Their Visualisation. *Sociology*, *52*(4), 830–848. https://doi.org/10.1177/0038038516674675

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. Cognitive Science, 11, 65–99.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. Medical Decision Making, 21, 37–44.

MacDonald-Ross, M. (1977). Graphics in texts. Rev. Res. Educ. 5: 49–85.

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2015). Improving risk understanding across ability levels: Encouraging active processing with dynamic icon arrays. *Journal of experimental psychology. Applied*, *21*(2), 178–194. https://doi.org/10.1037/xap0000045

Peck, E., & Ayuso, S., & El-Etr, O., (2019). Data is Personal: Attitudes and Perceptions of Data Visualization in Rural Pennsylvania. CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 1-12. 10.1145/3290605.3300474.

Peters, E., Vastfjall, D., Slovic, P., Mertz, C. K., Mozzocco, K., & Dickert, S. (2006). Numeracy and decision making. Psychological Science, 17, 406–413.

Petrova, D., & Garcia-Retamero, R., & Cokely, E., (2015). Understanding the Harms and Benefits of Cancer Screening: A Model of Factors That Shape Informed Decision Making. Medical decision making: an international journal of the Society for Medical Decision Making. 35. 10.1177/0272989X15587676.

Petty R.E., Cacioppo J.T. (1986) The Elaboration Likelihood Model of Persuasion. In: Communication and Persuasion. Springer Series in Social Psychology. Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4964-1_1

Petty, R. E., & Wegener, D. T. (1998). Matching versus mismatching attitude functions: Implications for scrutiny of persuasive messages. Personality and Social Psychology Bulletin, 24(3), 227–240.

Petty, R. E., Wegener, D. T., & Fabrigar, L. R. (1997). Attitudes and attitude change. Annual Review of Psychology, 48, 609–647.

Pinker S. A theory of graph comprehension. Pp. 73–126 in Freedle R (ed). Artificial Intelligence and the Future of Testing. Hillsdale, NJ: Erlbaum, 1990.

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*(1), 89–107. https://doi-org.ru.idm.oclc.org/10.1016/j.lindif.2007.03.011

Freedman, E., & Shah, P., & Vekiri, I., (2013). Shah, P., Freedman, E. G., & Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In P. Shah, & A. Miyake (Eds.), Cambridge Handbook of Visuospatial Thinking (pp. 426-476). New York, NY: Cambridge University Press..

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. *Manag. Sci., 66,* 2213-2231.

Stone, E. R., Sieck, W. R., Bull, B. E., Yates, J. F., Parks, S. C., & Rush, C. J. (2003). Foreground: Background salience: Explaining the effects of graphical displays on risk avoidance. Organizational Behavior and Human Decision Processes, 90, 19 –36. http://dx.doi.org/10.1016/ S0749-5978(03)00003-7

Tufte E.R., (2001) The Visual Display of Quantitative Information (Graphics Press LLC, Cheshire, CT).

Zacks, J., Levy, E., Tversky, B., & Schiano, D. (2002). Graphs in print. In M. Anderson, B. Meyer, & P. Olivier (Eds.), Diagrammatic representation and reasoning (pp. 187–206). London: Springer-Verlag.

Zacks, J., & Tversky, B. (1999). Bars and lines: A study of graphic communication. Memory & Cognition, 27, 1073–1079.

Zikmund-Fisher, B. J., Ubel, P. A., Smith, D. M., Derry, H. A., McClure, J. B., Stark, A., et al. (2008). Communicating side effect risks in a tamoxifen prophylaxis decision aid: the debiasing influence of pictographs. Patient Education and Counseling, 73, 209–214.

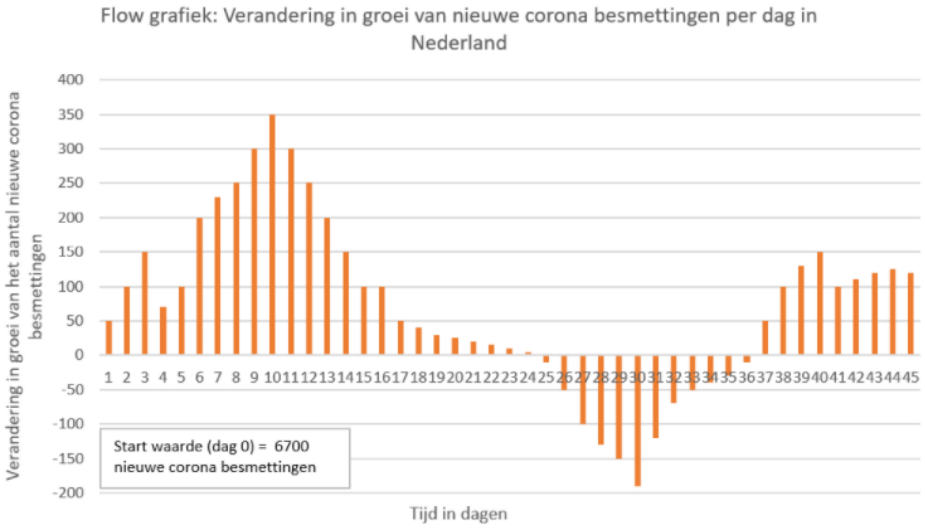
Appendix A

Questions measuring understanding:

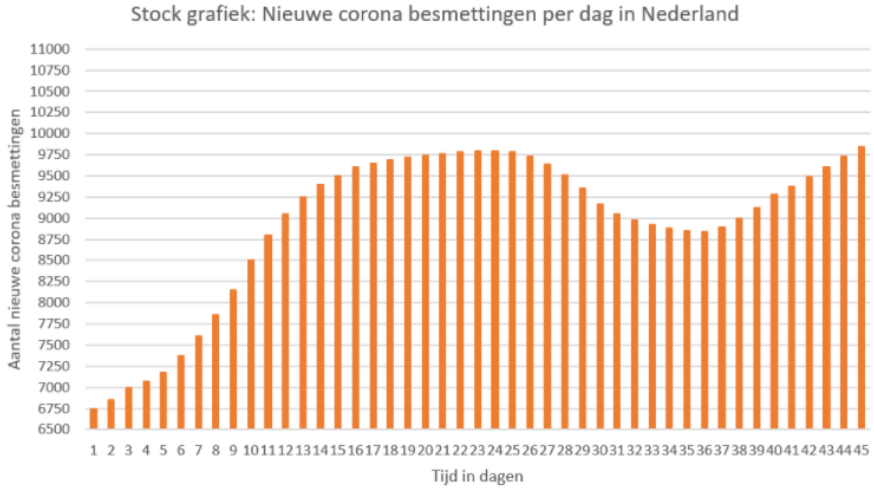
1. English: “How many new covid cases were registered on day 3?”
   Dutch: “Hoeveel nieuwe gevallen van COVID-19 waren er op dag 3?”
2. English: “Look at day 3 and day 4. Which day shows the biggest increase in infection compared to the day before?”
   Dutch: “Kijk naar dag 3 en dag 4. Welke dag laat de grootste stijging van nieuwe besmettingen zien ten opzichte van de dag ervoor?
3. English: “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?”
   Dutch: “Vergelijk de periode van dag 1 tot dag 10 met de periode van dag 11 tot dag 20. Welke periode laat een sterkere stijging in de besmettingen zien?”
4. English: “What is the difference in the number of new infections between day 12 and day 13?”
   Dutch: “Wat is het verschil in het aantal nieuwe besmettingen tussen dag 12 en dag 13?”
5. English: “On which day did the decrease of covid-19 infections start to slow down?”
   Dutch: “Op welke dag begon de snelheid van de daling in het aantal Covid-19 besmettingen af te nemen?”

Appendix B

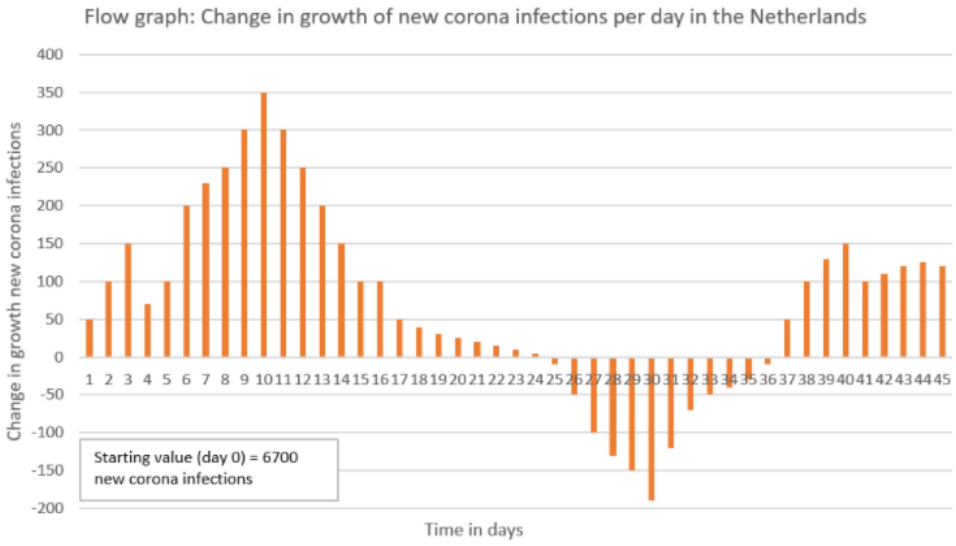
Stimuli used in the experiment in the analysed conditions:



Flow graph in high relevance condition, Dutch version

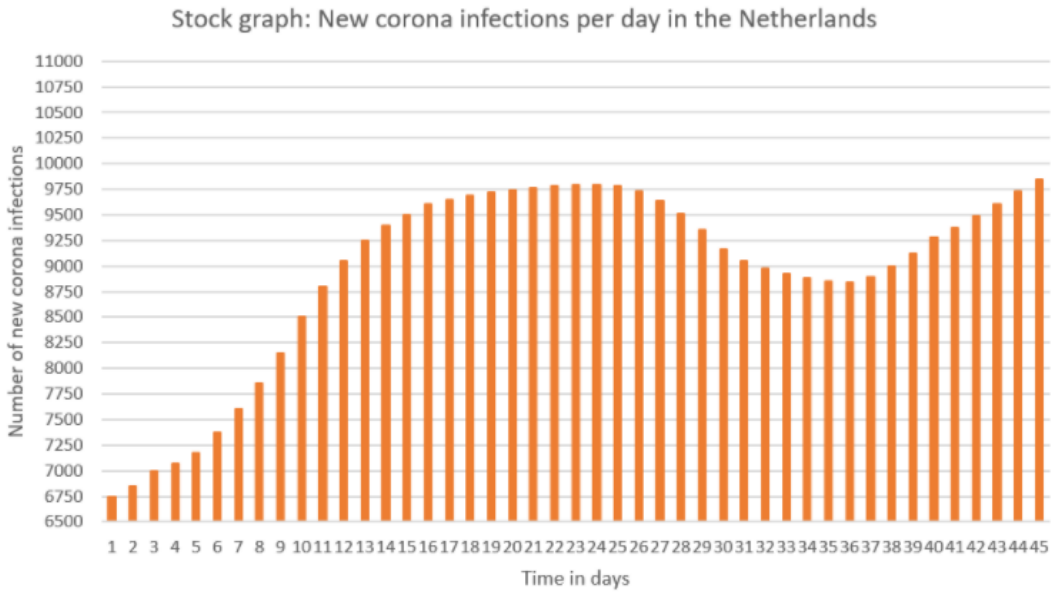


Stock graph in high relevance condition, Dutch version

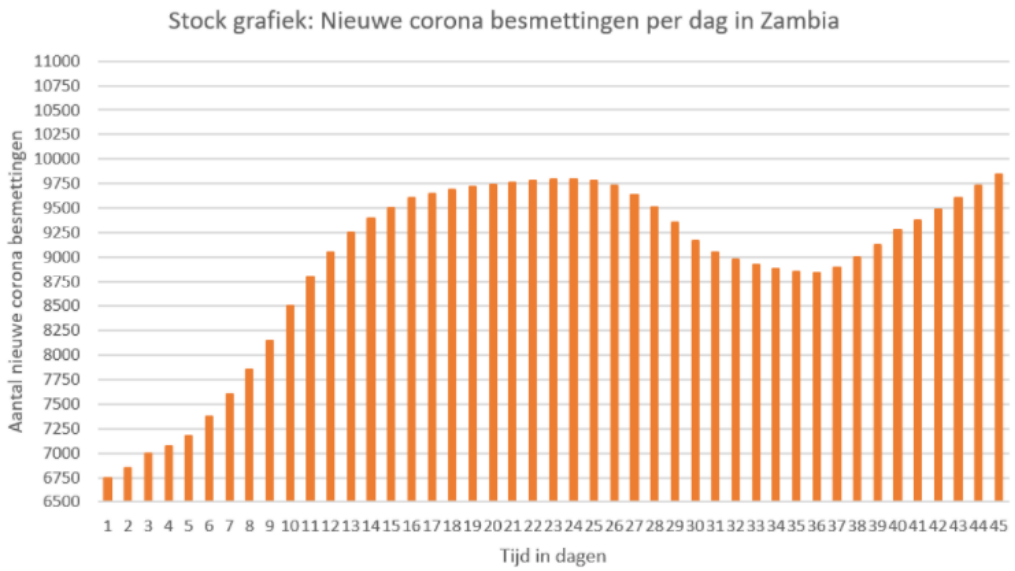


Flow graph

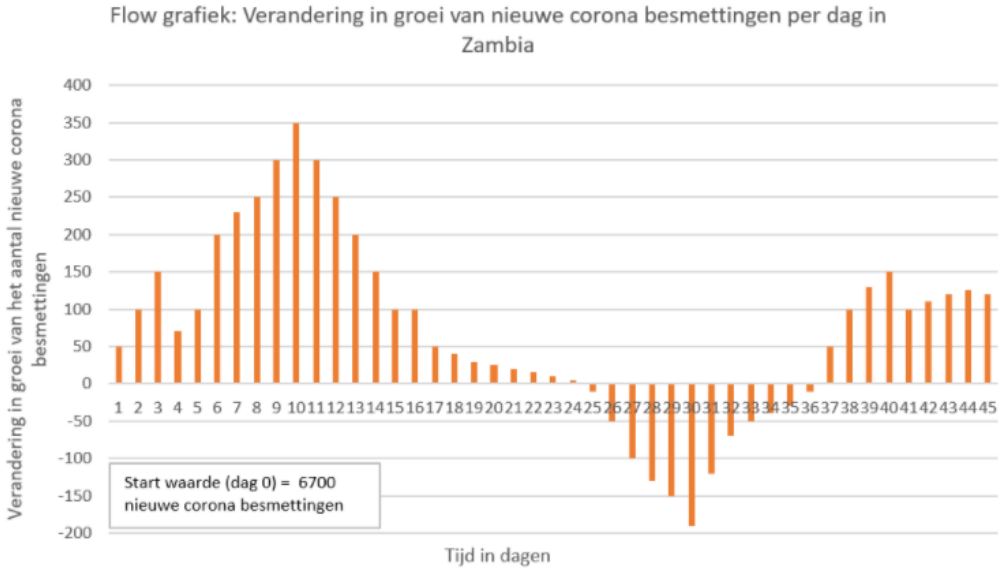
in high relevance condition, English version



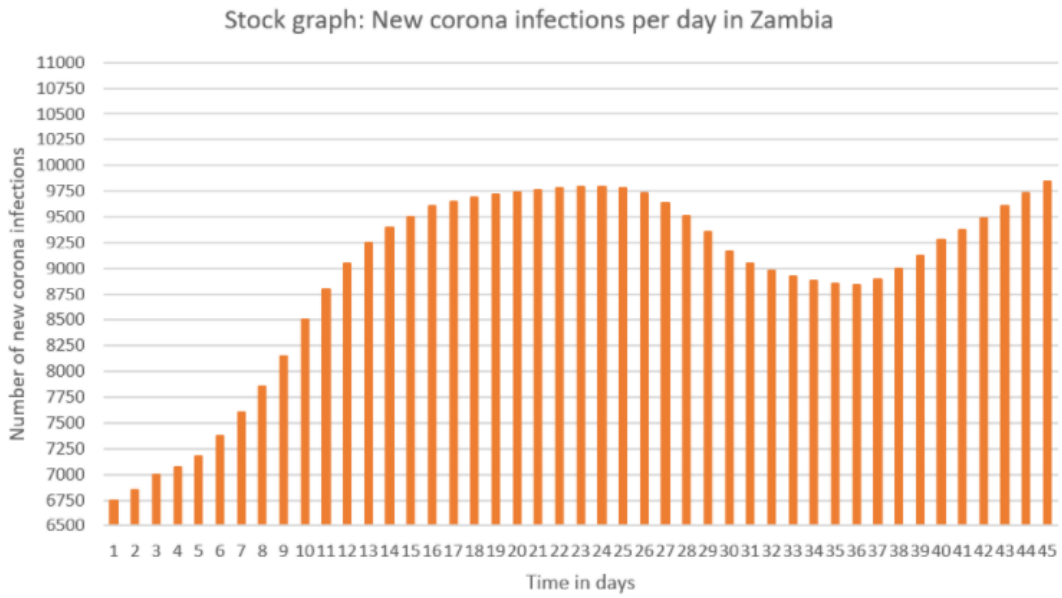
Stock graph high relevance condition, English version



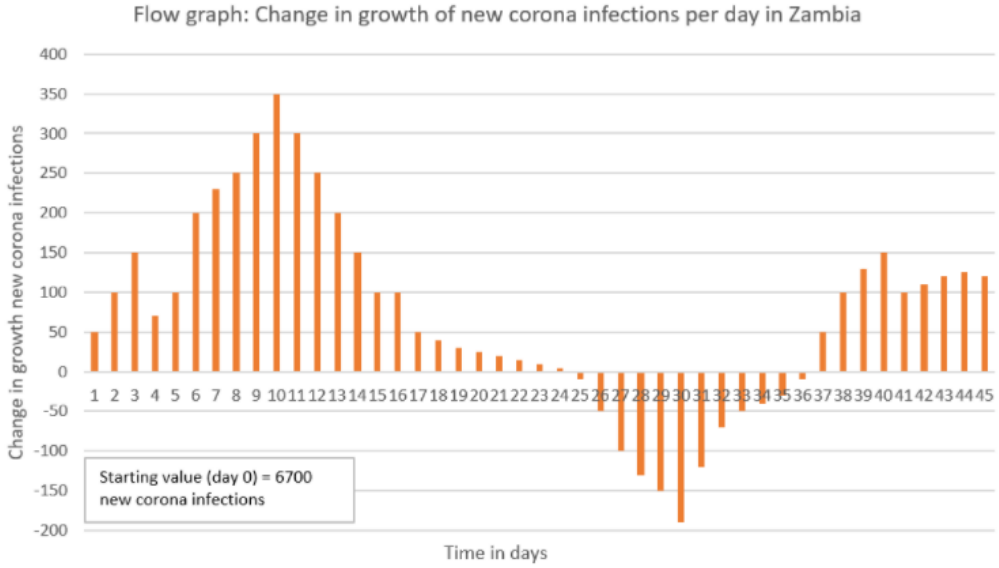
Stock graph low relevance condition, Dutch version



Flow graph low relevance condition, Dutch version



Stock graph low relevance condition, English version



Flow graph low relevance condition, English version

Appendix C

Statement of own work

Sign this Statement of own work form and add it as the last appendix in the final version of the Bachelor's thesis that is submitted as to the first supervisor.

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: Eindhoven 28-05-2021

Checklist ethics

Checklist EACH (version 1.6, november 2020)

You fill in the questions by clicking on the square next to the chosen answer ☐

After clicking, a cross will appear in this square ☒

1. Is a health care institution involved in the research?

Explanation: A health care institution is involved if one of the following (A/B/C) is the case:

A. One or more employees of a health care institution is/are involved in the research as principle or in the carrying out or execution of the research.
B. The research takes place within the walls of the health care institution and should, following the nature of the research, generally not be carried out outside the institution.
C. Patients / clients of the health care institution participate in the research (in the form of treatment).

☒ No → continue with questionnaire

☐ Yes → Did a Dutch Medical Institutional Review Board (MIRB) decide that the Wet Medisch Onderzoek (Medical Research Involving Human Subjects Act) is not applicable?

☐ Yes → continue with questionnaire

☐ No → This application should be reviewed by a Medical Institutional Review Board, for example, the Dutch CMO Regio Arnhem Nijmegen → end of checklist

2. Do grant providers wish the protocol to be assessed by a recognised MIRB?

☒ No → continue with questionnaire

☐ Yes → This application should be reviewed by a Medical Institutional Review Board, for example, the Dutch CMO Regio Arnhem Nijmegen → end of checklist

3. Does the research include medical-scientific research that might carry risks for the participant?☒ No → continue with questionnaire

☐ Yes → This application should be reviewed by a Medical Institutional Review Board, for example, the Dutch CMO Regio Arnhem Nijmegen → end of checklist

Standard research method

4. Does this research fall under one of the stated standard research methods of the Faculty of Arts or the Faculty of Philosophy, Theology and Religious Studies?

☒ Yes → ..... **(Standard questionnaire research, 12)** → continue with questionnaire

☐ No → assessment necessary, end of checklist

Participants

5. Is the participant population a healthy one?

☒ Yes → continue with questionnaire

☐ No → assessment necessary, end of checklist → go to assessment procedure

6. Will the research be conducted amongst minors (<16 years of age) or amongst (legally) incapable persons?

☐ Yes → assessment necessary, end of checklist → go to assessment procedure

☒ No → continue with questionnaire

Method

7. Is a method used that makes it possible to produce a coincidental finding that the participant should be informed of?

☐ Yes → assessment necessary, end of checklist → go to assessment procedure

☒ No → continue with questionnaire

8. Will participants undergo treatment or are they asked to perform certain behaviours that can lead to discomfort?

☐ Yes → assessment necessary, end of checklist → go to assessment procedure

☒ No → continue with questionnaire

9. Are the estimated risks connected to the research minimal?

☐ No → assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

10. Are the participants offered a different compensation than the usual one?

☐ Yes → assessment necessary, end of checklist → go to assessment procedure

☒ No → continue with questionnaire

11. Should deception take place, does the procedure meet the standard requirements?

☐ No → assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

12. Are the standard regulations regarding anonymity and privacy met?

☐ No → assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

Conducting the research

13. Will the research be carried out at an external location (such as a school, hospital)?

☒ No → continue with questionnaire

☐ Yes→ Do you have/will you receive written permission from this institution?

☐ No → assessment necessary, end of checklist → go to assessment procedure

☐ Yes → continue with questionnaire

14. Is there a contact person to whom participants can turn to with questions regarding the research and are they informed of this?

☐ No → assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

15. Is it clear for participants where they can file complaints with regard to participating in the research and how these complaints will be dealt with?

☐ No→ assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

16. Are the participants free to participate in the research, and to stop at any given point, whenever and for whatever reason they should wish to do so?

☐ No→ assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

17. Before participating, are participants informed by means of an information document about the aim, nature and risks and objections of the study? (zie explanation on informed consent and sample documents).

☐ No→ assessment necessary, end of checklist → go to assessment procedure

☒ Yes → continue with questionnaire

18. Do participants and/or their representatives sign a consent form? (zie explanation on informed consent and sample documents.

☐ No→ assessment necessary, end of checklist → go to assessment procedure

☒ Yes → checklist finished

If you want to record the results of this checklist, please save the completed file.

If you need approval from the EACH due to the requirement of a publisher or research grant provider, you will have to follow the formal assessment procedure of the EACH.