Data visualization & COVID-19

The influence of graph appearances on the emotion of the reader

Version 1

Date: 30-06-2021

University of Twente, Enschede

Author:	Merel Ensink (s2003015)	**Study program:**	Management, society & Technology
First supervisor:	Pieter-Jan Klok	**Word count:**	11942
Second supervisor:	Guus Dix	**Ethical approval:**	210480

1. Abstract

This research focuses on the research question whether graph appearances can influence the emotion of the reader regarding the COVID-19 situation in the Netherlands. To find a sufficient answer to this question, several hypotheses are tested to find a relationship between the appearances and the emotions. The tested appearances are *color, scale* and *type of number*. The tested emotions are *confusion, fear* and helplessness. The hypotheses are tested by analyzing data from the survey. In this survey, 100 respondents indicated to what extent their emotions were influenced after seeing different graphs. Almost all means for the variables were higher than 3, meaning that in general an emotion increased. It was concluded that all discussed appearances have an influence on the emotions of the reader. Darker colors resulted in more *fear.* The use of cumulative numbers resulted in more *helplessness* among the reader. For the use of a logarithmic scale, no significant effect was found for all three emotions. Furthermore, a significant relationship was found between the use of relative numbers and *confusion.* This can be explained by the fact that this graph is a bit harder to understand than the graphs with other types of numbers.

2. Table of contents

3. Background information 1
4. Theoretical framework 3
- **4.1** Emotions 3
- **4.2** Appearances 3
- **4.3** Hypotheses 4

5. Methods 6
- **5.1** Research design 6
- **5.2** Case selection and sampling 6
- **5.3** Operationalization 7

6. Data visualization on COVID-19 in the Netherlands 8
- **6.1** Channels 8
- **6.2** Visualization of corona data 8
- **6.3** Appearances 10
- **6.4** Conclusion 13

7. The influence of graph appearances on emotions 14
- **7.1** Sample and population 14
- **7.2** Descriptives indication question 16
- **7.3** Descritpives of the variables 16
- **7.4** Hypotheses 17
- **7.5** Conclusion 20

8. Conclusion and discussion 21
- **8.1** Answer to research question 21
- **8.2** Limitations 22
- **8.3** Recommendations 22

9. References 23
Appendix 1 – Survey 25
Appendix 2 – Graphs in the corona report 38
Appendix 3 – Graphs on coronavirus dashboard 42
Appendix 4 – Figures on subpage 51
Appendix 5 – Calculation gender 53
Appendix 6 – Distribution age 54
Appendix 7 – Descriptives for the variables 55

3. Background information

On Thursday the 27th of February 2020, the first COVID-19 infection in the Netherlands was reported (Rijksoverheid, 2020). Soon, measures such as the lockdown followed to limit the number of infections. Later, extra measures such as the curfew followed (Rijksoverheid, 2020).

Within this COVID-19 crisis, a lot of information is shared with citizens. This is mainly done through the use of graphs. Think about, for example, graphs with information about the number of infections, the number of hospital admissions, and the number of deceased persons. Within these types of graphs, there is also a difference in, for example, the use of relative, cumulative or absolute numbers. So, there is a diverse range of information that is shared with citizens via different sorts of graphs.

Furthermore, there are also differences in the acceptance of the measures and the assessment of the risks regarding the situation among the citizens of the Netherlands. The National Institute for Public Health and Environment (RIVM) together with all the different Municipal Health Services (GGD) have conducted several studies over time into, among other things, the rules of conduct that apply to the COVID-19 crisis in the Netherlands (RIVM, 2021). By doing these studies, the government hoped to be able to map out why people (do not) comply with the rules. Possible explanations mentioned here for (not) following the measures include risk assessment, emotions regarding the COVID-19 virus (e.g., fear, stress, helplessness), and the assessment whether the measures help (RIVM, 2021).

Some of these explanations for (not) following the measures can possibly be influenced by the way information is shared with the citizens. For example, it is possible that the emotion fear raises when the number of infected persons increases. However, it is also possible that this happens when cumulative numbers are used instead of relative numbers.

The aim of this research is to find out to what extent different graph appearances can influence the emotion of the reader. In this case, the reader is a citizen of the Netherlands. It is necessary to note that this research focuses solely on information shared by the government via official channels. Sharing this information is necessary in order to be as transparent as possible. When no information is shared about the situation, citizens will have no clear source that explains why certain measures are taken. When all information available is shared via the official channels of the government, it will be very hard for citizens to make sense of all this information. Finding out how graph appearances can influence the emotions of the reader can help with gaining knowledge about how information should be shared with the citizens. Especially in situations where, for example, the number of infected persons increases. To decrease this number, citizens should follow the measures taken by the government. If the appearance of a graph can influence how people feel about the COVID-19 situation and therefore can influence if they are willing to follow the measures, it can be a good tool for the government to protect the safety of citizens in terms of health. This shows the societal relevance of this research.

In terms of the scientific relevance, some information can be found on data visualization and even about emotions in engagement with data visualizations. However, most of the information that can be found about this focuses on emotions that arise from not understanding the particular visualization. For example, the article of Kennedy and Hill (2018), which states: “emotions like frustration and confusion sometimes resulted from participants’ consideration of whether the data had been visualized well or not” (Kennedy & Hill, 2018, p. 837). However, this article does not provide much information about which specific appearances of a graph can influence the emotions of a reader. Since little to no information can be found on this specific topic, it shows the scientific relevance of the current research.

So, not everyone follows the measures that are taken by the government in the COVID-19 situation, and this can be a risk for the public health. Because emotions can affect whether someone follows the measures and these emotions can potentially be affected by the graph appearances, the following research question was formulated:

"To what extent does the appearance of a graph influence the emotions of the citizens of the Netherlands regarding the COVID-19 situation?"

In order to get a sufficient answer to the research questions, the following sub-questions are established:

1. How is different information regarding the COVID-19 situation in the Netherlands shared by the government with the use of graphs?
2. How do characteristics of graph appearances influence the emotions of the reader of the graph regarding the COVID-19 situation?

"Emotions" is a rather broad principle. In the next section, the theoretical framework, these emotions will be discussed and demarcated.

4. Theoretical framework

In this section, emotions that can occur from reading graphs will be defined and discussed. Based on scientific articles, these emotions are linked to whether someone tends to comply with the measures. After that, options for appearances of graphs will be given based on available scientific articles and books. Based on this information, the hypotheses that will be tested in this research will be presented. First, a definition of data visualization should be mentioned. In this research, the definition by Andy Kirk in his book *Data visualization, a Handbook for Data Driven Design* will be used. This is: "the representation of data to facilitate understanding" (Kirk, 2016, p.31). In the specific case of the COVID-19 situation, this means that graphs are used to make the situation understandable for the citizens.

4.1 Emotions

In terms of emotions that can occur when reading a graph, we start with *confusion*. As mentioned before, frustration can occur when a graph is too hard to understand (Kennedy & Hill, 2018, p.837). This means that the reader is confused and will therefore get frustrated. This can in the end result in more risk-seeking actions instead of risk-averse actions (Hameleers, 2021). However, confusion does not only arise when a graph is too hard to understand, but also by misinformation. It is possible that this happens when a graph is made based on manipulated data (Goreis & Kothgassner, 2020). Thus, it is for example possible that confusion arises from the use of a wrong scale. Confusion does not only result in frustration, but also in less motivation. Previous research has shown that individuals that are feeling confused tend to give up earlier (Silvia, 2010, p.78-79). So, it is assumed that this also will result in more risk-seeking actions.

Second, the emotion of *fear*. Fear can be an important factor for compliance with measures taken by the government. Especially in pandemics, fear can result in health-compliant behavioral change, such as washing hands or social distancing (Harper et al., 2020). The emotion of fear in the COVID-19 situation can refer to several things. For example, the fear of getting the virus yourself or the fear of infecting someone else. However, it is possible that fear over time evolves into the feeling of being destined to be infected. This means that there is the feeling that an individual does not have any influence on the situation. When this is the case, individuals tend to ignore the preventive measures taken by the government (Kowalczuk & Gębski, 2021). The feeling that you don't have any influence on the situation can be defined as helplessness. However, while fear can result in helplessness, it is decided to separate these emotions in this research. This, since it is assumed that fear in general will have a positive effect on compliance with the measures, while helplessness will have a negative effect on compliance with the measures.

As mentioned above, the last emotion that will be discussed is *helplessness*. The emotion of helplessness can for example arise from media attention and reports on mortality (Dubey et al., 2020). Individuals that are feeling helpless will have the idea that they cannot do anything to change the situation. This also means that they do not believe that compliance with the measures will make any difference. However, in heavy cases, learned helplessness arises from for example a pandemic. Learned helplessness occurs when an individual is in a stressful situation repeatedly (Leonard, 2019). When (learned) helplessness occurs, individuals are known to do behavioral coping (Overmier, 2002). A very important factor for behavioral coping in learned helplessness is the environment of the person. This, since behavioral coping mostly happens in the personal network of a person. In the COVID-19 situation, this could result in more compliance with the measures by the government, however it is dependent on the environment of the person. This makes it a bit harder to say something about the relationship between helplessness and compliance with the rules. However, it is assumed that helplessness will have a negative effect on compliance with the measures.

4.2 Appearances

Now the emotions that can occur are explained, the different appearances of graphs can be discussed. One of the most important appearances that can possibly influence the emotions of a reader is the *color*. Using color can help in making a graph more readable, because color can be used to distinguish different categories in a graph. In terms of quantitative (ratio) data – for example the number of infected people – the data has an absolute zero point. Color can then help making a clear distinguish within this data

(Kirk, 2016, p.333). For example, the darker the color, the higher the number of infected people. Furthermore, the use of the color gray can help highlighting other parts of a graph (Kirk, 2016, p.341). More general information on the use of color can be found in the article by Clarke & Costall (2008), in which they focus on emotions that occur from different colors (Clarke & Costall, 2008). While it should be noted that personal experience and culture can have an influence on the rise of these emotions, some general information can be derived from this research (Clarke & Costall, 2008, p.409). First of all, cool colors such as green and blue seem to have a positive influence on the emotion of the reader, namely most participants in the research described this as relaxing and peaceful. On the other hand, warm colors such as red, yellow and orange provoked a whole range of emotions among the participants, ranging from happiness to anger (Clarke & Costall, 2008, p.407). Furthermore, the lighter the color, the more positive the emotions were for the participants. The other way around, the darker the color, the more negative emotions were mentioned by the participants. It even appeared to be the case that not the hue, but the saturation had the most influence on the emotions of about half of the participants (Clarke & Costall, 2008, p.408).

Second, a very important appearance of a graph is the *scale* that is used. When deciding on the scale, a decision should be made on the minimum and maximum value, but also on reasonable intervals (Kirk, 2016, p.364). The scale should be annotated in the graph, making it clear for the reader. Setting a higher or lower scale can have a significant influence on how the graph is interpreted by the reader. For example, when a graph is about percentages, but the scale is close to the value shown in the graph instead of 100%, the value can be interpreted as way bigger (Emery, 2017). Furthermore, having a very high maximum for a relatively low observed number can result in a wrong assessment of the situation. In the specific case of the COVID-19 situation, it could possibly cause that the situation is judged as less risky than it actually is. Furthermore, the New York Times published an article about the use of logarithmic scales instead of linear scales to visualize the exponential growth better in the COVID-19 situation. Where linear scales follow a linear interval on the y-axis, logarithmic scales have an increase on the y-axis regarding the previous number. When not constrained, the coronavirus spreads exponential. When there is a use of linear scales, the line of the curve shoots up. When a logarithmic scale is used, the curve is more a straight line (New York Times, 2020). However, it is argued that using a logarithmic scale makes the situation more confusion, especially for the broader audience (Romano et al., 2020). A research done by Ryan & Evers (2020), found that using a logarithmic scale resulted in several implications for the reader. Namely, miss assessment of the situation, perceiving the situation as less dangerous and expressing less support for the measures taken by the government and less intention of following these measures (Ryan & Evers, 2020). Thus, in terms of the emotions discussed before, the emotions fear and helplessness will probably be lower when using a logarithmic scale, while the emotion confusion will be greater.

Third, there can be made a distinction between what *type of numbers or percentages* are shown. In this research, we make a distinction between the use of absolute, relative or cumulative numbers. Some information can be found on the use of absolute numbers and risk assessment. Namely, the use of only absolute numbers can increase risk perception and therefore risk-averse behavior (Ancker et al. 2006). This, since it is a result from the increase of the emotion fear. The use of relative numbers seems to have an effect on confusion among the reader. Namely, the use of relative numbers, especially when there is no clear explanation, can result in more confusion among the reader (Baron, 1997). Little to no information can be found on the influence of cumulative numbers. However, it is assumed that this can have an influence on the emotion of the reader of the graph. It is for example possible that when a graph shows how many people died relative to the number of people infected, this affects how helpless the reader is feeling. This, because it can seem like the situation is getting worse and worse, even though it does not necessarily have to be the case.

4.3 Hypotheses

Based on this theory section, several hypotheses are formulated and tested. Possible other relations between the discussed appearances and emotions will also be tested. However, these are not formulated as hypotheses since the expected relationship is unclear.

4.3.1 Color

H1.1: The greater the difference of colors in the graphs, the lower the presence of the emotion confusion.
H1.2: The lighter the colors used in the graphs, the lower the presence of the emotion fear.
H1.3: The lighter the colors used in the graphs, the lower the presence of the emotion helplessness.

4.3.2 Scale

H2.1: The use of a logarithmic scale will result in a higher presence of the emotion confusion.
H2.2: The use of a logarithmic scale will result in a lower presence of the emotion fear.
H2.3: The use of a logarithmic scale will result in a lower presence of the emotion helplessness.
H2.4: The larger the range of the scale, the higher the presence of the emotion confusion.
H2.5: The larger the range of the scale, the lower the presence of the emotion fear.

4.3.3 Type of numbers

H3.1: The use of relative numbers will result in a higher presence of the emotion confusion.
H3.2: The use of absolute numbers will result in a higher presence of the emotion fear.
H3.3: The use of cumulative numbers will result in a higher presence of the emotion helplessness.

5. Methods

In this section, the methods for the research will be discussed. First, there will be deliberated on the research design. Then, information will be given on the case selection and sampling. Last, the operationalization of this research will be explained.

5.1 Research design

The research question of this research is a causal question. The causal research question posed in this research has the independent variable *appearances of a grap*h and the dependent variable *the emotion regarding the COVID-19 situation among the citizens of the Netherlands*. The units of analysis in this case are the citizens of the Netherlands. The research question posed is a causal question. This means that the focus is on finding the causal relationship between different variables. In this case, it focuses on the causal relationship between graph appearances and the emotion of a reader. The research design to answer this question is a causal design. There are three conditions that should be met to determine causality. First of all, there should be a correlation between the dependent and independent variable. Thus, there should be an association. Furthermore, the time order should be correct and there should be no intervening third variable. This also shows the possible threats of this research design. First, there is the threat of reverse causation. This means that it is possible that the dependent variable actually influences the independent variable, instead of the other way around. This would mean that the consequence is actually the cause of the cause. This seems like a legitimate threat, since it is possible that people interpret graphs differently due to the emotion regarding the COVID-19 situation that they already have. To limit this threat, the survey will start with the participant giving the current level they are feeling regarding the emotions that will be discussed. After that, they will have to indicate how the different graphs change these emotions (e.g., more confused). The second threat is that it is possible that there are third variables that influence the relation. Therefore, some control variables will be added to the survey. Such variables are for example age, gender or education level. The survey will be discussed more elaborate in the operationalization, chapter 5.3.

There are several steps to get a sufficient answer to the research question. First, graphs that are distributed by the Dutch government will be discussed in order to find out what information they perceive as necessary to help citizens understand the situation. Furthermore, it will help with developing the graphs for the survey. This, since it shows what visualizations of corona data are shared by the government. After that, the answers of the survey will be analyzed in order to find out to what extent characteristics of appearances of graphs can influence the emotion of the reader regarding the COVID-19 situation.

5.2 Case selection and sampling

Since this research focuses on the citizens of the Netherlands, the population is the same. However, to be triggered by data visualization, a person has to be able to at least slightly make sense of the visualization. Therefore, you could argue that the population is all the citizens of the Netherlands that are at least 16 years old. To gain information from this population, an online survey will be distributed via Facebook and Instagram. The goal is to have many cases among different age groups and education levels. To reach this goal, there will be made use of convenience sampling, a type of non-probability sampling (Edgar & Manz, 2017, p. 106). In this specific type of sampling technique, samples are picked because they are convenient located. Since this survey will be distributed via Facebook and Instagram, participating will be open for every citizen of the Netherlands that is 16 years or older and uses one of these platforms. However, it should be noted that this survey will be distributed via my personal accounts, meaning that the first participants will probably be my friends on these platforms. This can have a potential bias for the research since we come from the same environment and therefore these participants also will have about the same control variables. However, it will also be asked if people want to share the survey via their account. This creates a snowball effect that will result in more and other participants than only my own social media friends. Since the population is quite large, the sample size should also be quite large in order to draw a good conclusion. However, it should be noted that time is limited, so it might be hard to find a lot of participants within this time period. Some existing literature argues that a causal research should at least have a sample size of 50 participants (Okul et al., 2002). Others argue that for a large population the sample size should be at least 100 (Bullen, 2014). Therefore, it is decided to aim at a sample size of 100, but a sample size of at least 50 is perceived as sufficient for this research.

In terms of the platforms, studies on the daily use of social media platforms in the Netherlands in 2019 showed that Facebook is the biggest in the Netherlands with 7.1 million daily users of 15 years and older (Vader, 2020). According to the same study, Facebook is the most used social media platform amongst all age categories, except for the age category 15 - 19 years old. In this age category, Instagram is the most used platform. Furthermore, a research of I&O research in 2017 showed that there were no significant differences in the level of education among the users of Facebook and Instagram and that all levels of education are present on both platforms (I&O research, 2020). Therefore, the survey will be distributed on both Facebook and Instagram to get the most representative sample. However, it should be noted that the distribution of the survey via social media is a limitation for the research, since it does not include the part of the population that does not use social media.

5.3 Operationalization

Most of the data needed for answering the research question will come from the survey. Since this research focuses on the emotions of the citizens of the Netherlands after reading graphs, the survey was distributed in Dutch. However, the English translation of the survey is attached as appendix 1. As can be seen in the appendix, the survey starts with a short explanation of the research, followed by questions about the demographic features of the participant. After that, the definitions of the different emotions are given. This is done, because it is important that all the participants use the same definition for answering the questions. The definitions are followed by a question where the participants must indicate on a scale of 1 - 10 to what extent the emotions are present before seeing any graph. As mentioned before, this is done in order to prevent reverse causation. In the rest of the survey, the participants get to see different graphs for every appearance mentioned in the theory section. For every appearance, the graphs are based on the same data. This is done to prevent that the participant is influenced by the data instead of the appearance. Above the graph, the participant gets to see the answers on the first question. So, they see what their initial emotions were. This is done because the participant needs to answer to what extent their emotions are influenced by the graph as a comparison to their initial feeling. They do that by indicating on a scale from 1 – 5 to what extent their emotions changed. Here, 1 means that the feeling is greatly reduced, 3 means that the feeling is unchanged and 5 means that the feeling is greatly increased. This can be answered in a matrix with the following statement above: “after seeing this graph, my emotions are:”. In the matrix, the three emotions are displayed in the row and the scale from 1 – 5 is displayed in the column.

For the appearance *type of number or percentage,* three graphs are developed. For the other appearances, *scale* and *color*, four graphs each are developed. However, each participant only gets to see two graphs for these two appearances. This is divided by a randomizer. As mentioned before, the graphs are based on the same data. The graphs with the appearances *scale* and *color* look similar, which makes it easy for the participant to see that the situation has not changed. Therefore, it is decided to show only two graphs for these two appearances. The rest of this section is devoted to giving some information on how the data will be analyzed.

For analyzing the data, partly filled in answer sheets will in general be deleted. This prevents double answers by participants that stopped before finishing and started later on a new answer sheet. However, this can be prevented by checking the IP address, provided that the respondent uses the same device. Therefore, when a survey is completed for more than three quarters and the IP address does not appear more often, it will be included in the analysis. After the partly filled in answer sheets are deleted, a one sample t-test will be performed. The data will be tested with a value of 3, which meant that an emotion was unchanged. Therefore, the null hypothesis for the analysis is also that a feeling was unchanged after seeing a graph.

The next section focuses on answering the first sub-question. After that, the data of the survey will be analyzed in order to give a sufficient answer to the second sub-question and the research question.

6. Data visualization on COVID-19 in the Netherlands

How is different information regarding the COVID-19 situation in the Netherlands shared by the government with the use of graphs?

As mentioned before, this research will solely be based on information shared via official channels of the government of the Netherlands. Several characteristics are studied in order to give a sufficient answer to the first sub-question. In this section, it will be discussed via which channels the government shares information, what visualizations of corona data are shared and to what extent the different appearances can be observed.

6.1 Channels

First of all, the channels used by the government. While the press conferences are an important tool for sharing information with the citizens, there is no use of data visualization there. However, there are several websites where graphs regarding the COVID-19 situation in the Netherlands can be found. The official websites used by the government for sharing information are the one of the RIVM and the website "Coronavirus Dashboard". The latter is mostly promoted by the government as channel for sharing information.

On the website of the RIVM, various information can be found. First of all, there is a link to a list of several data-and information sources regarding the COVID-19 situation. The reason why this link is provided is in order to make it easier for the citizens of the Netherlands to find information, but also to ensure that the citizens are not informed by fake news. In addition to the list of sources, there is also a link to a dataset. This dataset is the one that is used by the RIVM for the development of the graphs, and it is open and available for everyone. The actual graphs made by the RIVM can be found on two places of the websites. First of all, there is the subpage "development COVID-19 in graphs" (RIVM, 2021). Here, four different weekly updated graphs are displayed. Next to this subpage, there are the reports that are called "weekly update of the epidemiological situation of COVID-19 in the Netherlands" (RIVM, 2021). These reports are all open and available. However, these reports do use some jargon and are somewhat difficult to understand for the broad public. Furthermore, the reports consist of about 80 pages and 50 figures. It takes some time to read through this report and to make sense of the data. Therefore, it is assumed that the subpage is used more often by the broad public than the reports. The graphs on both the subpage and in the report are either based on data coming from the RIVM, GGD or the National Intensive Care Evaluation Foundation (NICE).

The website of the coronavirus dashboard shares several types of information and is daily updated. This is done by using graphs, but also by sharing news articles. Furthermore, a distinction is made for the whole country, per municipality or per safety region. The website also states that it aims to be openly available for everyone on every device. It is emphasized that the developers of the website aim at making the information as clear as possible, so it is understandable for everyone. The website is available in Dutch and English, since some residents of the Netherlands do not speak Dutch. Last, there is a page on the website for the justification of the shared information. Here, the different headings that are used for the visualization of corona data are displayed. Under each of these headings, an explanation is given of the source that is used, on which numbers it is based and how the visualization should be interpreted (Rijksoverheid, 2021).

There are more webpages of the government that share information about the corona situation. However, these pages are either fully or mostly fully filled with text instead of visualizations. So, some graphs can be found here. However, it decided to leave these pages out, since these graphs generally come from one of the websites mentioned before.

6.2 Visualization of corona data

Second, the visualizations of corona data that are shared by the government. As mentioned before, there are several visualizations that are communicated with the citizens. On the subpage, there is a graph for the reported positively tested persons per day, a graph on the number of hospital admissions per day, a graph on the number of deceased persons a day and a graph on the number of deceased persons by age and gender (RIVM, 2021).

For the report, the different visualizations of corona data can be found in appendix 2. It should be noted that this appendix is made on the basis of the report published on the 27th of April. However, the graphs used in the other weekly reports have the same purpose. Furthermore, not all graphs are added to the appendix. This, since the same graphs were made starting from 2020 as for 2021. Except for the timespan, the appearances are the same. Therefore, it is decided to leave them out. As can be seen in the appendix, most of the graphs focus on the same visualizations of corona data as the first subpages, namely: the number of positive tests, number of deceased persons, and the number of hospital admissions. However, the report has way more extensive graphs than the subpage. For example, they make a distinction per age category and per safety region. Furthermore, there is a focus on nursing homes, residential care centers and disabled care institutions (RIVM, 2021). Additionally, there are also graphs on the estimated number of contagious persons and the reproduction number R. The latter refers to how many people on average are infected by one patient with COVID-19 (RIVM, 2020). Also, there are graphs that combine the number of tests with the percentage positive tests. One interesting thing done in the report is adding several graphs into one figure. For example, figure 1. As can be seen, the graphs of hospital admissions and intensive care admissions are put in the same figure. This allows the reader to see how the situation develops and how severe the cases are over time. On the other hand, ti can confuse the reader since the graphs use a different scale on the y-ax. However, since these are two different graphs with different information they are treated as such in the appendix.

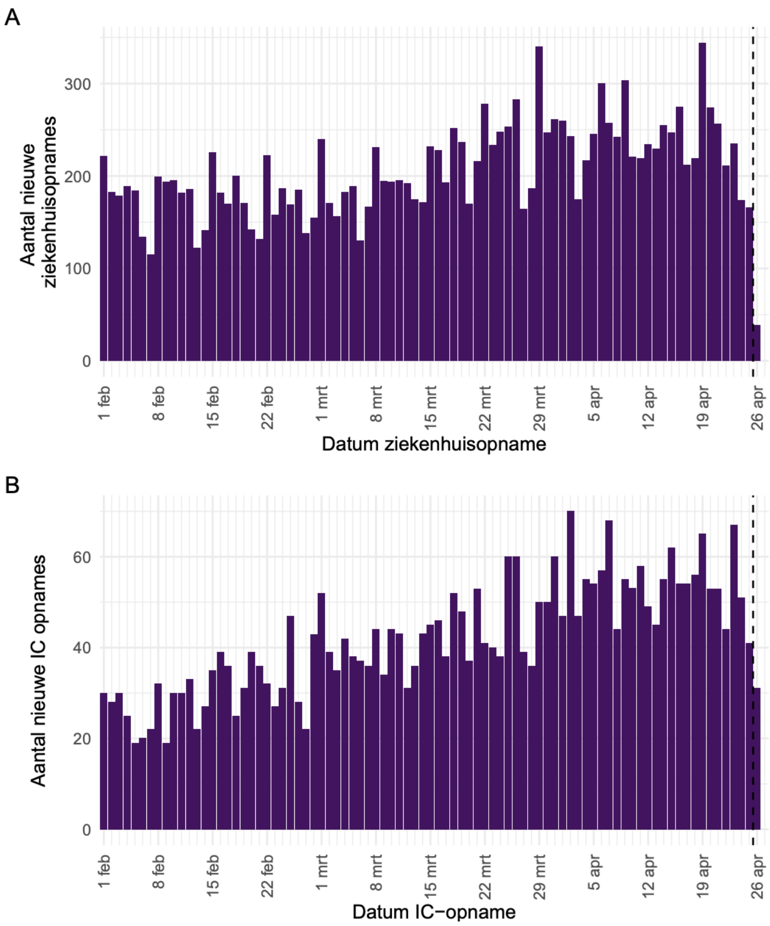


Figure 1: Number of hospital admissions per day, where A is hospital admission and B is Intensive Care admission

For the coronavirus dashboard, information about the different visualizations of corona data can be found in appendix 3. This is based on graphs on the 14th of June. However, the graphs are daily updated and have the same purpose. Since the website is updated daily, it is possible that certain appearances differ. It is for example possible that over time the scale changes, because of a reduction in infections. Furthermore, it should be noted that not all graphs are added to the appendix. As mentioned before, the coronavirus dashboard has separate graphs for the whole country, per municipality or per safety region. Since these graphs have the same appearances, but focus on different parts of the data, it is decided to focus only on the graphs for the whole country. As can be seen in appendix 3, the different visualizations are divided under the headings: *vaccinations*, *hospitals*, *infections*, *behavior*, *vulnerable groups*, *early signal*s and *other*. Every heading consists of at least one graph. A lot of graphs are focusing on the same visualization of corona data as the subpage and the report. Thus, the number of infections, the number of hospital admission and so on. However, one new graph is the one on the number of vaccinations per day. Furthermore, the corona dashboard also focuses on sewage measurements and data coming from the corona detector app (Rijksoverheid, 2021). This app gives a notification when you have been close to someone who later turns out to have corona for at least 15 minutes (Rijksoverheid, 2021).

6.3 Appearances

This last section is devoted to finding out to what extent appearances discussed in the theoretical framework can be found in the graphs shared by the government. For the subpage, the graphs can be found in appendix 4. As mentioned before, for both the report and the coronavirus dashboard, a summary of the appearances per graph can be found in appendix 2 and 3. However, some general information will be given and some noteworthy graphs will be discussed.

6.3.1 Subpage

As mentioned above, the graphs of the subpage can be found in appendix 4. The first graph on the subpage, figure 9, is the one of the positively tested persons per day. Here, they use the color purple for the positive tests that were already on the website before the weekly update and the color yellow for the new ones. Furthermore, the scale on the y-axis has a minimum of zero and a maximum of 15.000 with an interval of 5.000. The types of numbers that are used are absolute numbers. The second graph, figure 10 is the one of the numbers of hospital admissions per day. These are shown with the color purple. The y-axis has a minimum of zero and a maximum of 400 with an interval of 100. Here, the types of numbers are again absolute. The third graph, figure 11, is the one of the numbers of deceased persons per day also uses the color purple and yellow, just as in the first graph. However, as can be seen in, there are still deceased persons added to, for example, the end of April, while this weekly update is done mid-June. Therefore, you can conclude that for some people the cause of death, COVID-19, is added later. The y-axis has a minimum of zero and a maximum of 40 with an interval of 20. In terms of numbers, there is also the use of absolute numbers. The last graph, figure 12, is the one on the number of deceased persons by age and gender. Here, they use the color blue for men and dark pink for women. On the x-axis, the age categories are displayed consisting of 4 years per category. The y-axis has a minimum of zero and a maximum of 6000 with an interval of 2000. Here, there is also the use of absolute numbers.

6.3.2 Report

In terms of color, most colors that are used have a high saturation. Only in cases where there is uncertainty if the numbers are correct, lighter colors are used. This is for example done in figure 2, where the visualization is about the number of deceased persons by age and gender. Since there is no duty to report when a person dies from COVID-19, there is underreporting. Sometimes it also appeared that colors mentioned in the legenda were very hard to find in the graph. This is for example the case for the gender type "not mentioned", which also can be found in figure 2. This type can be found on the x-axis, but is hard to see what the value is for this gender type. Another noteworthy fact is that for almost all graphs a different color is used to separate the new data from the old data. Only the graph for the number of hospital admissions and intensive care admissions, figure 1, uses a dotted line to indicate the new data.

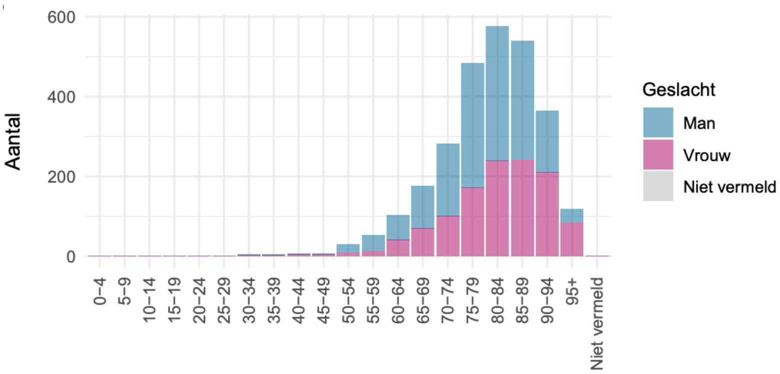


Figure 2: number of deceased persons per day by age and gender

In terms of the scale, sometimes the y-axis continues a bit after the last displayed number, such as in figure 3. The maximum value for the scale mentioned in appendix 2 is the maximum displayed number on the y-axis. So, in figure 3 the y-axis goes beyond 20, but 20 is the maximum value mentioned in the appendix. Almost all graphs have a minimum of zero, except for the graph on total mortality per week compared to the expected mortality. There is a logical explanation for this, namely based on previous years it is noted that the mortality will not be lower than 2000 per month. Lastly, it should be noted that all scales are linear, just as the graphs on the subpage.

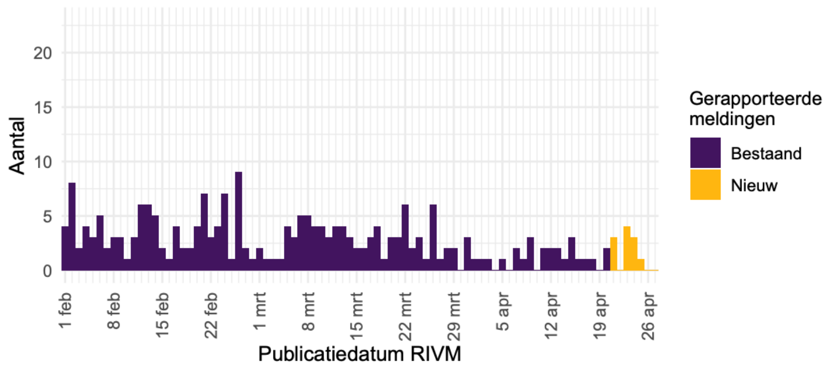


Figure 3: Number of disabled care institutions with at least one infection

In terms of the types of numbers, almost every graph used absolute numbers. Sometimes percentages were used. For example in figure 4, where the percentage positive tests out of all the tests is visualized. Here, the percentage positive tests is a relative number. No cumulative numbers were used in the graphs. In one case, the one of the reproduction number, it is neither absolute, relative or cumulative. This, since the type of number is just the reproduction number.

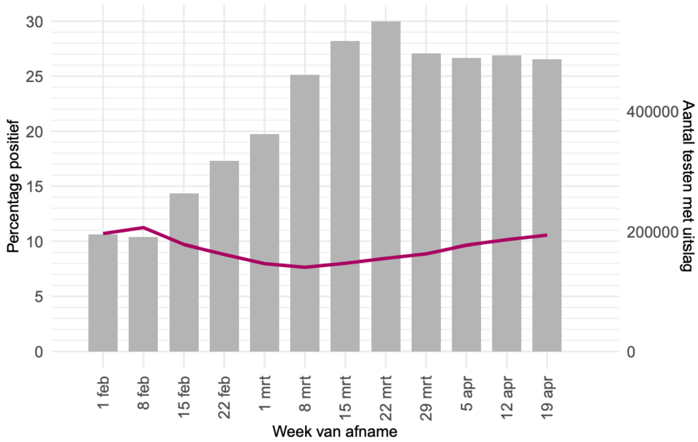


Figure 4: Total positive tests and percentage positive tests per week

6.3.3 Corona dashboard

In terms of color, most graphs on the coronavirus dashboard use a blue hue. In some graphs, different types of blue are used. For example, in figure 5 about the number of tests over time. There are also graphs that use a lot of different colors. For example, figure 6, about the hospital admissions per age category. This figure is a bit harder to understand, because of all the different colors. However, on the website there is also the possibility to select only the age category or categories that you are interested in. Furthermore, in figure 6 the last days are gray, because they are not complete. Gray is used in multiple graphs on the dashboard when it is incomplete or uncertain. However, this is not a rule of thumb. Figure 5, for example, deviates from this, as gray is used for tests with a result.

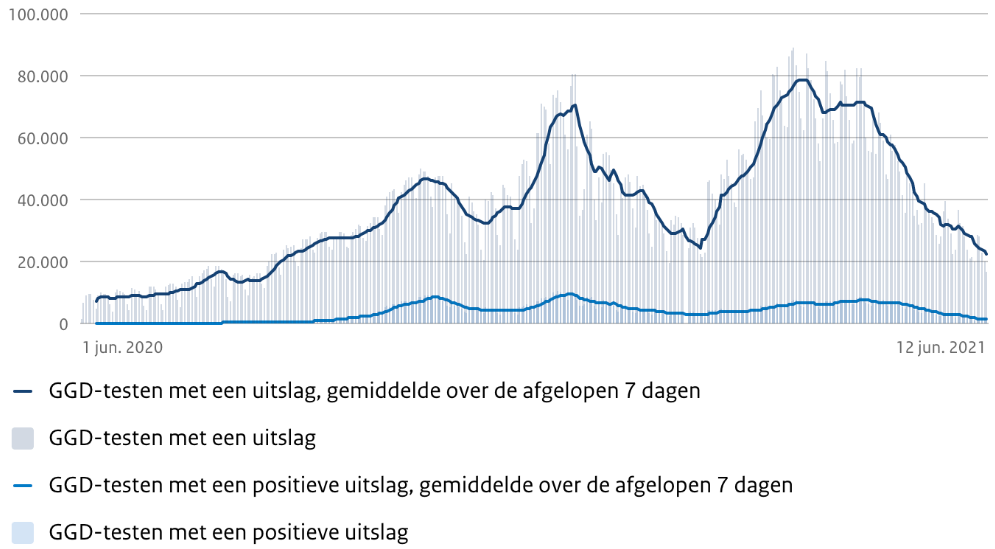


Figure 5: number of tests over time

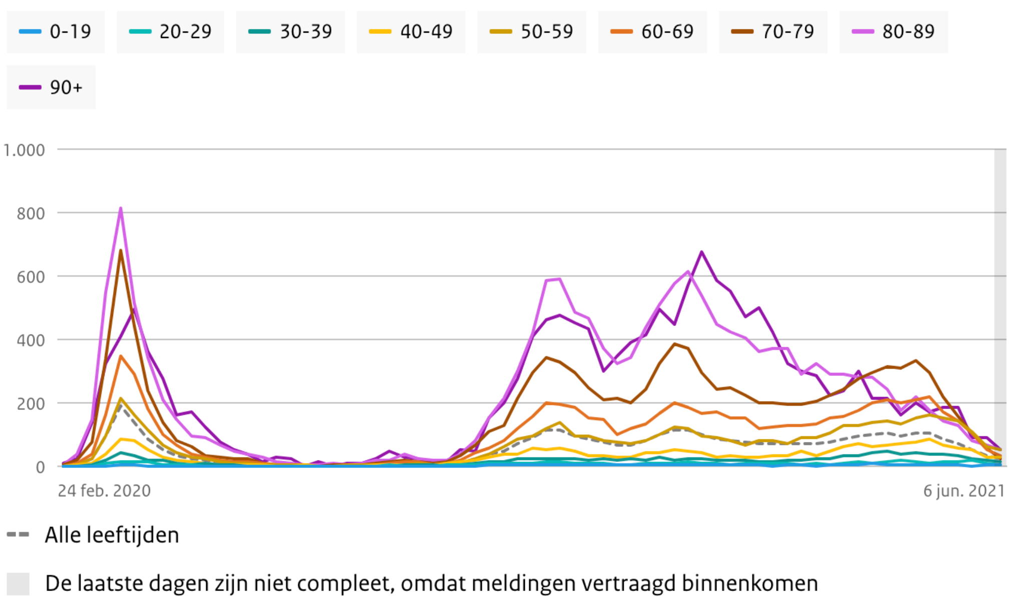


Figure 6: number of hospital admissions over time per age category

All graphs on the coronavirus dashboard use a linear scale. Furthermore, all graphs have a minimum of zero. However, the graph on the growth rate, figure 7, goes down to a maximum of -40. This is the only graph with negative numbers on the y-axis. This means that there were less people tested positively on corona than on average the 7 days before. One other noteworthy thing is that the graphs on virus particles in wastewater use a scale ranging from 0 to 800. However, the actual number should be calculated by multiplying it with 100 billion.

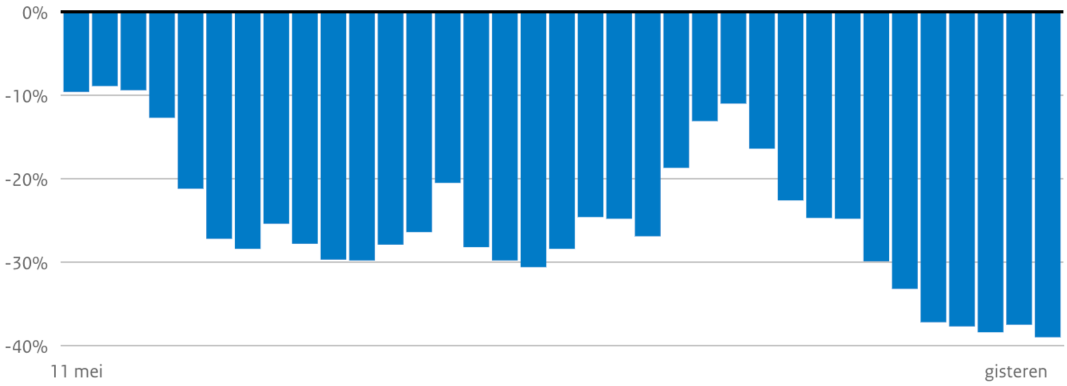


Figure 7: growth rate over time

Most graphs on the dashboard use absolute numbers. However, there are some graphs that use relative numbers. For some graphs, this is quite clear. Such as for figure 7, where the percentages are even on the y-axis. However, figure 6 for example also uses relative numbers. The number displayed on the graph, is the number of hospital admissions of a certain age category, per 1.000.000 persons in this age category. There is also one graph on the dashboard that uses cumulative numbers. This graph, figure 8, focuses on the total number of vaccinations. Furthermore, there is also a graph on the reproduction number on the dashboard, meaning that this graph neither uses absolute, relative or cumulative numbers.

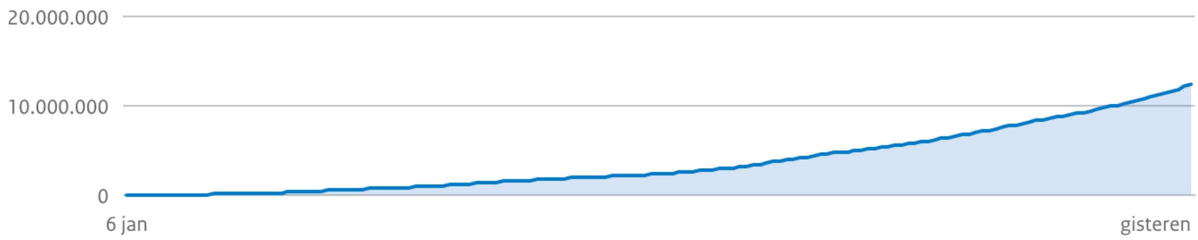


Figure 8: total number of vaccinations

6.4 Conclusion

In this section, the first sub-question of the research is answered. First, the channels via which the government shares it graphs were discussed. This was via the website of the RIVM, its weekly reports and the coronavirus dashboard. Second, the different visualizations of corona data that are shared were discussed. While the report and the coronavirus dashboard were much more extensive than the subpage, they all had similarities in the visualizations of corona data. This is because they treat the same situations (e.g. hospital admissions). Where the subpage only shows one graph for this, the report and the dashboard share several. Last, it is discussed to what extent the different appearances discussed in the theoretical framework can be linked to the graphs. In terms of the color, the subpage does not make a distinction in terms of the lightness of the color. The report does have a difference in lightness for situations where the data uncertain. The coronavirus dashboard mostly uses the color blue, but sometimes uses a lot of colors to make a distinction between several categories. In terms of scale, all the graphs use a linear scale and almost all graphs have a minimum of zero. On the coronavirus dashboard, there is one graph where the number on the y-axis decreases instead of increases. In terms of the types of numbers, on the subpage it were all absolute numbers. In the report, most graphs used absolute numbers, but sometimes relative numbers were used. On the coronavirus dashboard, there were absolute, relative and cumulative numbers. However, on the dashboard it was not always immediately clear that it were relative numbers. As mentioned before, this is the case in figure 6. Both the report and the coronavirus dashboard also had a graph on the reproduction number. This graph neither uses absolute, relative or cumulative numbers.

7. The influence of graph appearances on emotions

“How do characteristics of graph appearances influence the emotions of the reader of the graph regarding the COVID-19 situation?”

In this section, the data from the survey will be analyzed to give an answer to the second sub-question. First, something will be said about the sample and it will be compared to the population with the use of the control data. Then, the variables will be described and the hypotheses will be tested.

7.1 Sample and population

In total, 152 respondents started the survey. 100 of these respondents did finish the survey. The other 52 respondents filled in between 3 and 55 percent of the survey. As mentioned before, it was decided to only use surveys that were filled in for at least three quarters. This means that in this research, N = 100. Now, this sample should be compared to the population to see whether it is representative. This will be done by using the variables gender, age and education.

7.1.1 Gender

As can be seen in table 1, gender was distributed not entirely even in the sample. 41% of the respondents were men, 57% percent were women and 2% did identify as else. To compare it with the population, data from the Central Bureau of Statistics (CBS) will be used. Since there is only data available from the 1st of January 2020, this will be used. According to the CBS, in general there are more women than men in the Netherlands. This is also the case in this sample. The distribution of gender according to the CBS is for every 100 women, there are approximately 99 men (CBS, 2020). Since the CBS does not include a category “else”, these are left out of the sample for the comparison calculation. This calculation can be found in appendix 5. In this research, the sample would be representative when there would be a bit more than 49 women and a bit less than 49 men. This is not the case in this research, which means that in terms of gender, the sample is not completely representative, as man are under-represented.

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Else	2	2,0	2,0	2,0
	Man	41	41,0	41,0	43,0
	Woman	57	57,0	57,0	100,0
	Total	100	100,0	100,0	

Table 1: distribution of gender in the response

7.1.2 Age

In terms of the age category, the distribution in the sample is quite uneven. The youngest respondent was 19 and the oldest 76. As can be seen in table 2, 87% of the respondents have an age ranging from 19 to 44. Figure 13 in appendix 6 gives a clearer visualization of the distribution of age in the sample. Here, it can also be seen that the mean age in this sample is 27,29. Once again, the sample will be compared to data from the CBS. According to this data, in 2020 the mean age in the Netherlands was 42.2 years (CBS, 2020). According to the same data, 25% of the citizens of the Netherlands in 2020 was in the age category 20 to 40 years old. In the sample, this is 81%. So, it can be concluded that the sample is not representative to the population. This, since the difference between the sample and the population is very large. However, it should be noted that this difference is the biggest in the group of 20 – 24, as this category is 65% of the respondents.

Age_categories

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	< 20	6	6,0	6,0	6,0
	20-24	65	65,0	65,0	71,0
	25-29	12	12,0	12,0	83,0
	30-34	2	2,0	2,0	85,0
	35-39	2	2,0	2,0	87,0
	40-44	0	0,0	0,0	87,0
	45-49	3	3,0	3,0	90,0
	50-54	2	2,0	2,0	92,0
	55-59	4	4,0	4,0	96,0
	60-64	0	0,0	0,0	96,0
	65-69	2	2,0	2,0	98,0
	70-74	1	1,0	1,0	99,0
	> 74	1	1,0	1,0	100,0
	Total	100	100,0	100,0	

Table 2: frequency of age in the response

7.1.3 Education

The last variable that needs to be checked is education. In the survey, the respondent had to indicate which level of education they were following, or which level of education they had followed. The distribution of the education level of the sample can be found in table 3. As can be seen, most of the respondents either follow or followed higher vocational education, or education on university-level. When comparing this to the data from the CBS, it can be concluded that this also does not correspond with the population. According to this data, 41% of the population has either followed higher vocational education or university level education (CBS, 2020). In the sample, this is 90%, as can be seen in table 3. This shows that the sample is not representative for the population. This can cause a bias, since higher educated people will presumably be better able to understand the data visualization.

Education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Else, namely:	1	1,0	1,0	1,0
	Preparatory secondary vocational education	1	1,0	1,0	2,0
	Pre-university education	2	2,0	2,0	4,0
	Post-secondary vocational education	6	6,0	6,0	10,0
	Higher vocational education	30	30,0	30,0	40,0
	University (Bachelor)	29	29,0	29,0	69,0
	University (Master)	31	31,0	31,0	100,0
	Total	100	100,0	100,0	

Table 3: distribution of education in the response

7.1.4 Conclusion

From the comparison it should be concluded that the sample is not representative for the population of the Netherlands. In terms of gender, the difference is not that big. For age however, the sample is in general way younger than the population. This can influence the research, since younger people will be in general less sick from COVID-19 than the elderly. Therefore, the risk assessment and emotions toward the situation can be different. In terms of education level, there is also a significant difference. The sample is in general higher educated than the population. This can be explained by the distribution of the survey via my own social media network. This can influence the results of the research since higher educated people will probably be able to make more sense of the visualization. Therefore, it is possible that emotions of these readers are less influenced than would be the case in the actual population.

The next section will be devoted to explaining the descritpives of the variables and testing the hypotheses. However, it should be kept in mind that the sample of the research in general is younger and higher educated than the population is.

7.2 Descriptives indication question

As mentioned before, the survey started with asking the respondents to what extent the discussed emotions (confusion, fear and helplessness) were present before participating in the survey. The descriptives of the answers to this question can be found in table 4. For both *fear* and *confusion*, N=100 while for *helplessness* this N=99. This means that one respondent indicated that they rather did not say to what extent the emotion helplessness was present. All three variables have a minimum of 1, meaning that at least one respondent indicated that the emotion was not present. There was no respondent that indicated that the feeling of *confusion* was very present. Therefore, the maximum value is 9, while for the other emotions this is 10. As can be seen in the table, all means are below 5. The variable *fear* has the lowest mean (2,7) and the lowest standard deviation (1,6). The low standard deviation means that the given answers cluster around 2,7. For *confusion*, the mean is a bit higher (3,6) just as the standard deviation (2,3). For *helplessness* both the mean (4,1) and standard deviation (2,8) are the highest. These indication variables give an insight in how the respondents were feeling before seeing a graph. The rather low mean can be explained by the fact that the COVID-19 situation is going on for over a year. Since people got a lot of information to process in this period, which can result in people getting fed up with the situation. It also should be noted that when a respondent (for example) answered that a feeling was not present at all, that this feeling cannot be reduced.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Confusion_indication	100	1	9	3,62	2,308
Fear_indication	100	1	10	2,68	1,632
Helplessness_indication	99	1	10	4,10	2,819
Valid N (listwise)	99				

Table 4: Descriptives of the indication question

7.3 Descriptives of the variables

As mentioned before, all respondents had the questions about the type of number. As can be seen in appendix 7, this means that for the graph with absolute numbers, the graph with relative numbers and the graph with cumulative numbers N = 100. The other questions were randomized, meaning that not everybody had the same question. For the graphs that focused on *scale,* the data was divided in N=48 and N=52. For *color*, the questions were divided even, meaning that for both groups N=50. Furthermore, as can be seen in appendix 7, all variables have a mean close to 3 and no respondent answered with a value of 6. In the survey, 6 meant that a respondent did not want to tell if the emotion changed and 3 meant that the emotions of the reader were unchanged. So, there does not seem to be a demonstrable difference in the emotions after seeing a graph. However, this will be tested in the next section. A last thing that should be noted from appendix 7 is the skewness. In general, a skewness higher than -1 or 1 means that the distribution is highly skewed and therefore not normally distributed. A skewness of 0 means that the distribution is exactly normally distributed. Not all the variables are normally distributed, however since the sample size for all the variables is bigger than 30, this will not influence the results of the one sample t-test (van den Berg, 2021).

7.4 Hypotheses

In this section, the hypotheses will be tested. As mentioned before, all means are close to 3. However, if the difference in emotions is significant is investigated by using a one sample t-test in SPSS. This is done by testing it with a test value of 3. A value of 3 in the survey meant an emotion was unchanged after seeing a graph. So, the null hypothesis that is tested is that an emotion is unchanged after seeing a graph.

This section is divided in different sub-paragraphs for every independent variable. So, one for color, one for scale and one for the type of number. All results of the hypotheses per independent variable are merged into one table. When there is no significant effect, this is mentioned. More information is given for the hypotheses that have a significant effect.

7.4.1 Color

The results of the hypotheses for color can be found in table 5. An effect is significant when the p-value is <0.05. So, the hypotheses that have no significant effect are: **H1.2.1 (**p = 0,371), **H1.3.1** (p = 0,083) and **H1.3.2** (p = 0,128). The exact formulations of these hypotheses are:

H1.2.1: *The lighter the colors used in the graphs, the lower the presence of the emotion fear.*
H1.3.1: *The lighter the colors used in the graphs, the lower the presence of the emotion helplessness.*
H1.3.2: *The darker the colors used in the graphs, the higher the presence of the emotion helplessness.*

The other hypotheses did have a significant effect. So, more information of these hypotheses will be given below.

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
H1.1: Different colors confusion	2,365	49	,022	,260	,04	,48
H1.2.1: Light color fear	0,903	49	,371	,060	-,07	,19
H1.2.2: Dark color fear	2,850	49	,006	,320	,09	,55
H1.3.1: Light color helplessness	1,769	49	,083	,120	-,02	,26
H1.3.2: Dark color helplessness	1,549	49	,128	,140	-,04	,32

Table 5: one sample t-test for the hypotheses regarding color

H1.1: *The greater the difference of colors in the graphs, the lower the presence of the emotion confusion.* The observed mean for the emotion *confusion* after seeing a graph with different colors is 3,26 as can be seen in table 13. As can be seen in table 5, the p-value for this hypothesis is 0,022. Since this p-value is <0.05, the hypothesis that the emotion confusion is unchanged can be rejected, with t(49) = 2,37, p = 0.022. However, because of the higher mean it is more likely that the emotion confusion is increased than decreased. So, we cannot conclude that H1.1 is accepted, but we can conclude that using different colors has a significant effect on the emotion confusion of the reader.

H1.2.2: *The darker the colors used in the graphs, the higher the presence of the emotion fear*.
No significant effect was found on the emotion *fear* when light colors were used. However, it can be useful to know whether the use of darker colors has a significant effect. Therefore, this hypothesis is added. The mean answer for the emotion fear with dark colors in the graph is 3,32 as can be seen in table 14. The p-value is 0,006, meaning that $p<0,05$. Therefore, we can reject the hypothesis that the emotion *fear* for graphs with a dark color is unchanged, with $t(49) = 2,85$, $p = 0,006$. Since the mean value is higher than 3, the hypothesis that darker colors result in a higher presence of the emotion fear can be accepted.

7.4.2 Scale

The results of the test for the hypotheses that focus on scale can be found in table 6. From this table it can be concluded that there is no significant effect found in these hypotheses. This, because all p-values are above 0,05. **H2.1**, **H2.2** and **H2.5** even have a p-value of 1. These were also the variables that had a mean of exactly 3. A p-value of 1 means that it can be assumed that the result is founded by chance. So, for every variable tested here, the null hypothesis is accepted. This means that the emotion is unchanged after seeing the graph. The initial hypotheses were:

H2.1: *The use of a logarithmic scale will result in a higher presence of the emotion confusion.*
H2.2: *The use of a logarithmic scale will result in a lower presence of the emotion fear.*
H2.3: *The use of a logarithmic scale will result in a lower presence of the emotion helplessness.*
H2.4: *The larger the range of the scale, the higher the presence of the emotion confusion.*
H2.5: *The larger the range of the scale, the lower the presence of the emotion fear*.

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
H2.1: Logarithmic scale confusion	,000	47	1,000	,000	-,21	,21
H2.2: Logarithmic scale fear	,000	47	1,000	,000	-,20	,20
H2.3: Logarithmic scale helplessness	1,741	47	,088	,167	-,03	,36
H2.4: Big scale confusion	-1,399	51	,168	-,096	-,23	,04
H2.5: Big scale fear	,000	51	1,000	,000	-,15	,15

Table 6: one sample t-test for the hypotheses regarding scale

To test whether scale can have an influence on emotions at all, the same emotions are tested again, but with the opposite graphs. Thus, instead of a logarithmic scale, the emotions after seeing the graph with a linear scale are tested. Instead of a big scale the normal scale is tested. This means that the following hypothesis are established:

H2.6: *The use of a linear scale will result in a lower presence of the emotion confusion.*

H2.7: The use of a linear scale will result in a higher presence of the emotion fear
H2.8: The use of a linear scale will result in a higher presence of the emotion helplessness
H2.9: The use of a normal scale will result in a lower presence of the emotion confusion
H2.10: The use of a normal scale will result in higher presence of the emotion fear

As can be seen in table 7, **H2.9** is the only hypothesis that did not found a significant effect. Therefore, it is concluded that there is not enough evidence to reject the null hypothesis that the emotion *confusion* is unchanged after seeing a graph with a normal scale.

The hypotheses that focused on the emotions after seeing a linear scale all have a significant effect. Therefore, it can be concluded that there is enough evidence to reject the null hypothesis that the emotions *confusion, fear* and *helplessness* are unchanged after seeing a graph with a linear scale. Furthermore, they all have a mean higher than 3, as can be seen in appendix 7. This means that it is hard to say whether **H2.6** can be accepted, since the hypothesis is that the emotions of *confusion* would be lower. **H2.7** and **H2.8** can be accepted as the hypothesis states that the emotions *fear* and *helplessness* would be higher. When looking at the survey in appendix 1, this increase in the emotions can be explained. As can be seen, the logarithmic scale has a rather straight line, where the line for the linear scale increases a lot, especially in the end. This can cause more fear and helplessness for the reader.

H2.10 is accepted with t(47) = 2,441, p = 0,018. This means that there is enough evidence to say that a normal scale will result in a higher presence of the emotion *fear*. Since the data for developing the graphs are the same for the dependent variables, this can also be explained by looking at the survey in appendix 1. Namely, the number of deceased persons fluctuates, but increases in the end. This can cause more fear for the reader.

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
H2.6: Linear scale confusion	2,108	51	,040	,192	,01	,38
H2.7: Linear scale fear	2,431	51	,019	,173	,03	,32
H2.8: Linear scale helplessness	4,173	51	,000	,288	,15	,43
H2.9: Normal scale confusion	-,206	47	,837	-,021	-,22	,18
H2.10 Normal scale fear	2,441	47	,018	,188	,03	,34

Table 7: one sample t-test for the extra hypothesis regarding scale

7.4.3 Type of numbers

The results of the test for the hypotheses for type of numbers can be found in table 8. As can be seen, only hypothesis **H3.2** can be rejected. This, because the p-value 0,566>0,05. Therefore, we can conclude that the emotion *fear* is unchanged for a graph with absolute numbers. The exact formulation of this hypothesis was:

H3.2: The use of absolute numbers will result in a higher presence of the emotion fear.

H3.1: *The use of relative numbers will result in a higher presence of the emotion confusion.*
The observed mean of the emotion *confusion* for a graph with relative numbers was 3,2 as can be seen in table 20 As can be seen in table 7, the p-value is 0,016, which is >0,05. Therefore, the null hypothesis that the emotion *confusion* is unchanged for a graph with relative numbers can be rejected, with t(99) = 2,449, p = 0,016. Since the mean is higher than 3, it can also be concluded that the hypothesis that the use of relative numbers will result in a higher presence of the emotion confusion is accepted. When looking at appendix 1, a simple explanation can be found. Since the total amount of tests is displayed as a bar graph, with a line for the relative number of positive tests, the graph can be complex to understand. Therefore, readers of this graph can get confused.

H3.3: *The use of cumulative numbers will result in a higher presence of the emotion helplessness.*
As can be seen in table 21, the observed mean value for the emotion *helplessness* for graphs with cumulative numbers is 3,47. This is the second highest observed mean. As can be seen in table 7, the p-value is 0,000. This does not mean that this value is exactly 0, but it is rounded. Therefore, it can be concluded that the null hypothesis that the emotion *helplessness* is unchanged after seeing a graph with cumulative numbers is rejected. Since the observed mean is above 3, it can also be concluded that the hypothesis that the use of cumulative numbers will result in a higher presence of the emotion helplessness is accepted.

One-Sample Test

	Test Value = 3					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
H3.1: Relative numbers confusion	2,449	99	,016	,200	,04	,36
H3.2: Absolute numbers fear	,575	99	,566	,030	-,07	,13
H3.3: Cumulative numbers helplessness	7,915	99	,000	,470	,35	,59

Table 8: one sample t-test for the hypotheses regarding type of number

7.5 Conclusion
The tested hypotheses make it possible to give an answer to the second sub-question. There is a significant effect found for at least two hypotheses per independent variable, but not all hypotheses were equally significant. However, it can be concluded that graph appearances have an influence on the emotion of the reader. However, not every hypothesis that had a significant effect could be accepted. This, because **H1.1** and **H2.6** stated that the presence of the emotion *confusion* would be lower. However, since all observed means of the tested hypothesis were higher than 3, it is hard to accept the hypothesis that the presence of the emotion would be lower. It can be concluded that the use of darker colors will result in a higher presence of the emotion *fear* of the reader. Furthermore, no significant effect was found when using a logarithmic or a large scale. However, the opposite scales, linear and small, did found a significant effect. The use of relative numbers did have a significant effect on the emotion *confusion* and the use cumulative numbers did have a significant effect on the emotion *helplessness*.

One thing that should be noted, is that for the independent variables *scale* and *color* the sample size is about 50. This is rather small and therefore outliers can have an influence on the effect. However, it is hard to exclude outliers, since there is a scale of 1 – 5 and these values are therefore not so far apart.

8. Conclusion and discussion

This last section is devoted to giving an answer to the research question. Furthermore, the limitations for this research and recommendation for new research are discussed.

8.1 Answer to research question

In this research, the following research question is addressed:

"To what extent does the appearance of a graph influence the emotions of the citizens of the Netherlands regarding the COVID-19 situation?"

For graph appearances, a distinction was made between color, scale and type of numbers. For emotions, this was confusion, fear and helplessness. Furthermore, it was discussed that finding an answer to this research question could explain why people (do not) comply with measurements by the government. Namely, it is assumed that *fear* would result in more compliance with the rules, while *confusion* and *helplessness* would result in less compliance with the rules. To get a sufficient answer to the research question, the following sub-questions were addressed:

1. How is different information regarding the COVID-19 situation in the Netherlands shared by the government with the use of graphs?
2. How do characteristics of graph appearances influence the emotions of the reader of the graph regarding the COVID-19 situation?

In terms of the first sub-question, there are three official channels of the government discussed that were used for sharing information about the COVID-19 situation with the use of graphs. These were the subpage of the RIVM, the corona report and the coronavirus dashboard. The coronavirus dashboard is the channel that is mostly promoted by the government as a source for finding information. Both the report and the dashboard had a lot of graphs. The sub-page only had four. There were similarities in these graphs, as a lot of them focused on the same visualization of corona data (e.g. hospital admissions). However, the use of the appearances *color*, *scale* and *type of number*s differed among these channels. In terms of lightness of color, the report and the dashboard used a lighter color for uncertain situations. Furthermore, the most often used color on the dashboard was (light) blue and on the subpage and the report it was dark purple. Distinctions in categories for all the three channels were mostly made by using very different colors. All the graphs had a rather normal scale, meaning that the observed value fitted the scale. Furthermore, there were only linear scales, so no logarithmic ones. In terms of the types of numbers, the subpage only used absolute numbers. The report used for some graphs relative numbers, but it were mainly graphs with absolute numbers. The coronavirus dashboard used both absolute and relative numbers and even one graph with cumulative numbers.

The second sub-question focused on to what extent graphs can influence the emotions of the reader regarding the COVID-19 situation. The answers of the survey were analyzed to answer this question. Before respondents answered the question to what extent their emotions changed after seeing a graph, they were asked how they were feeling regarding the COVID-19 situation. It was concluded that the emotions were not highly present at the respondents. This can indicate that people are fed up with the situation and therefore do not really care anymore. For the graph appearances, it can be concluded that they have an influence on the emotion of the reader, as there is a significant effect found for at least two hypotheses per independent variable. All hypotheses that were accepted stated that a certain emotion increased when reading a particular graph. From these hypotheses, it can be concluded that the use of darker colors will result in a higher presence of the emotion *fear* of the reader. It can also be concluded that the use of different colors influences the emotion *confusion*. However, it cannot be concluded that it lowers the emotion. Furthermore, a significant effect on all three emotions was found when using a linear scale. This could be explained by the raising slope of the line. The normal scale also had a significant effect on the emotion *fear*. This could also be explained by the slope of the line. Last, the use of relative numbers did have a significant effect on the emotion *confusion* and the use cumulative numbers did have a significant effect on the emotion *helplessness*.

As mentioned in the theoretical framework, it is assumed that *fear* has a positive effect on compliance with rules, while *confusion* and *helplessness* have a negative effect. There is no evident effect found that resulted in a lower presence of *confusion* and *helplessness*. However, it was found that there was enough evidence to conclude that the use of darker colors and a linear scale resulted in a higher presence of the emotion *fear*. The answer to sub-question 1 already showed that the government does only uses a linear scale in their graphs. Furthermore, a lot of graphs on the subpage and in the report used a dark purple color. This can result in more *fear* amongst the citizens and, as mentioned in the theoretical framework, it is assumed that this will result in more compliance with rules. The report and the coronavirus dashboard both used graphs with relative numbers. The analysis of the data showed that the use of relative numbers resulted in a higher presence of the emotion *confusion*. In the theoretical framework, it was mentioned that the emotion *confusion* assumingly resulted in people less complying with the measures. Therefore, it can probably be helpful for the government to use less graphs with relative numbers or to provide an explanation on how to read the graph. This will assumingly result in a lower presence of the emotion *confusion*.

8.2 Limitations

There are several limitations of this research. First of all, the survey was distributed via my own social media network. As can be seen in the descriptives of the respondents, this had the consequence that the respondents were in general younger and higher educated than the population. Furthermore, since there was a randomizer for the variables scale and color, these had a rather small sample size, namely $N \approx 50$. Therefore, the analysis is more sensitive to outliers. Since time was limited, the survey had to close eventually to start the analysis. More time could have probably result in more respondents. Furthermore, an attempt was made to account for third variables. However, this did not work out completely. It is for example possible that people do not trust the government. Therefore, their emotions will probably be unchanged when seeing a graph that is based on data from the government. Furthermore, as mentioned in section 7, sometimes it is not clear if the graph appearance itself influence the emotion of the reader. It is possible that a respondent was influence by the slope of the line rather than the graphical appearance itself. A last limitation that should be noted is that this research focuses on the COVID-19 situation. As mentioned in the background, the first infection in the Netherlands was reported on the 27th of February 2020. This means that when the survey was distributed, the COVID-19 situation had been going on for over a year. In this year, people are overloaded with information about the situation. In conversations, some respondents indicated that they cared less or did not even care at all about the situation. Therefore, it is a limitation that this survey focuses on the COVID-19 situation.

8.3 Recommendations

Stemming from the last limitation a recommendation is made. Namely, it is recommended to examine again the relationship between graph appearances and emotions. However, for the next research it is recommended to focus on another situation than COVID-19. This, because people are overloaded with information about COVID-19, which makes them care less. Therefore, it is assumed that the emotions were less influenced than would be the case when the research focused on something else. It is also recommended to have a higher sample size for a survey in a next research.

The second recommendation focuses on the compliance with the measurements. In this research it was stated that emotions can have influence on whether someone follows the rules. However, some information available on this contradicts each other. This was the case for (learned) helplessness. Where some sources stated that this could result in more compliance with the rules, others stated that it would be the other way around. Therefore, it is recommended to do more research on the influence of emotions on the willingness to comply with the measures.

9. References

Ancker, J. S., Senathirajah, Y., Kukafka, R., & Starren, J. B. (2006). Design Features of Graphs in Health Risk Communication: A Systematic Review. *Journal of the American Medical Informatics Association*, *13*(6), 608–618. Retrieved from: https://doi.org/10.1197/jamia.M2115

Baron, J. Confusion of Relative and Absolute Risk in Valuation. *Journal of Risk and Uncertainty* 14, 301-309 (1997). Retrieved from: https://doi.org/10.1023/A:1007796310463

Bullen, P. B. (2014). *How to choose a sample size (for the statistically challenged)*. Retrieved from: http://www.tools4dev.org/resources/how-to-choose-a-sample-size/

CBS. (2020). *Leeftijdsverdeling.* Retrieved from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/leeftijd/bevolking

CBS. (2020). *Mannen en vrouwen.* Retrieved from: https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking/mannen-en-vrouwen

CBS. (2020). *Onderwijs.* Retrieved from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71450ned/table?fromstatweb

Clarke, T., & Costall, A. (2008). The emotional connotations of color: A qualitative investigation. *Color Research and Application*, *33*(5), 406–410. Retrieved from: https://doi.org/10.1002/col.20435

Dubey, S., Biswas, P., Ghosh, R., Chatterjee, S., Dubey, M. J., Chatterjee, S., Lahiri, D., & Lavie, C. J. (2020). Psychosocial impact of COVID-19. *Diabetes and Metabolic Syndrome: Clinical Research and Reviews*, *14*(5), 779–788. Retrieved from: https://doi.org/10.1016/j.dsx.2020.05.035

Edgar, T. W., & Manz, D. O. (2017). Exploratory Study. In *Research Methods for Cyber Security* (pp. 95–130). Elsevier. Retrieved from: https://doi.org/10.1016/b978-0-12-805349-2.00004-2

Emery, A. K. (2017). *The Graph's Scale: Actual Maximum Value or Potential Maximum Value?* Retrieved from: https://depictdatastudio.com/axis/

Goreis & Kothgassner (2020). Social Media as a Vehicle for Conspiracy Beliefs about COVID-19. *Digital Psychology.* Retrieved from: 10.24989/dp.v1i2.1866

Hameleers, M. (2021). Prospect Theory in Times of a Pandemic: The Effects of Gain versus Loss Framing on Risky Choices and Emotional Responses during the 2020 Coronavirus Outbreak–Evidence from the US and the Netherlands. *Mass Communication and Society*. Retrieved from: https://doi.org/10.1080/15205436.2020.1870144

Harper, C. A., Satchell, L. P., Fido, D., & Latzman, R. D. (2020). Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic. *International Journal of Mental Health and Addiction*. Retrieved from: https://doi.org/10.1007/s11469-020-00281-5

I&O research. (2020). *Laaggeletterdheid op de onderzoekskaart*. Retrieved from: https://www.binnenlandsbestuur.nl/bestuur-en-organisatie/kennispartners/i-o-research/laaggeletterden-minder-op-facebook.13281454.lynkx

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

Kowalczuk & Gebski (2021). Impact of Fear of Contracting COVID-19 and Complying with the Rules of Isolation on Nutritional Behaviors of Polish Adults. *International journal of Environmental Research and Public Health, (18)*4, 1631. Retrieved from: **https://doi.org/10.3390/ijerph18041631**

Kirk, A. (2016). *Data Visualisation: A Handbook for Data Driven Design*. London, UK: SAGE

Leonard, J. (2019). *What is learned helplessness?* Retrieved from: https://www.medicalnewstoday.com/articles/325355

New York Times. (2020). *A different way to chart the spread of the coronavirus.* Retrieved from: https://www.nytimes.com/2020/03/20/health/coronavirus-data-logarithm-chart.html

Okul, İ. /, Öğretmen, Ö., Bilimsel, A., Becerilerine, S., Alan, İ., & Delice, A. (2002). *The Sampling Issues in Quantitative Research*. Retrieved from:

Overmier, J. B. (2002). *On Learned Helplessness*. Retrieved from: https://link.springer.com/content/pdf/10.1007/BF02688801.pdf

Rijksoverheid (2020). *Avondklok.* Retrieved from: https://www.rijksoverheid.nl/onderwerpen/coronavirus-covid-19/avondklok

Rijksoverheid. (2020). *Coronamelder.* Retrieved from: https://www.coronamelder.nl/nl/

Rijksoverheid. (2020). *Current situation in the Netherlands.* Retrieved from: https://coronadashboard.government.nl/

Rijksoverheid. (2020). *Februari 2020: Eerste coronabesmetting in Nederland*. Retrieved from: https://www.rijksoverheid.nl/onderwerpen/coronavirus-tijdlijn/februari-2020-eerste-coronabesmetting-in-nederland

RIVM. (2020). *Hoe rekenmodellen bijdragen aan de bestrijding van COVID-19.* Retrieved from: https://www.rivm.nl/coronavirus-covid-19/rekenmodellen

RIVM. (2021). *Ontwikkeling COVID-19 in grafieken.* Retrieved from: https://www.rivm.nl/coronavirus-covid-19/grafieken

RIVM. (2021). *Resultaten onderzoek gedragsregels en welbevinden*. Retrieved from: https://www.rivm.nl/gedragsonderzoek/maatregelen-welbevinden

RIVM. (2021). *Verklaring gedrag*. Retrieved from: https://www.rivm.nl/gedragsonderzoek/maatregelen-welbevinden/verklaringen-gedrag

RIVM. (2021). *Wekelijkse update epidemiologische situatie van SARS-CoV-2 in Nederland.* Retrieved from: https://www.rivm.nl/coronavirus-covid-19/actueel/wekelijkse-update-epidemiologische-situatie-covid-19-in-nederland

Romano, A, Sotis, C, Dominioni, G, Guidi, S. The scale of COVID-19 graphs affects understanding, attitudes, and policy preferences. *Health Economics.* 2020; 29: 1482-1494. Retrieved from: https://doi.org/10.1002/hec.4143

Ryan, W., & Evers, E. R. K. (2020). Logarithmic Axis Graphs Distort Lay Judgment. Retrieved from: https://doi.org/10.31234/osf.io/cwt56

Silvia, P. (2010). Confusion and interest: The Role of Knowledge Emotions in Aesthetic Experience. *Psychology of Aesthetics Creativity and the Arts* 4(2): 75-60. Retrieved from: DOI:10.1037/a0017081

Vader, C. (2020). *Social Media in Nederland in 2020: Facebook stijgt licht, Instagram stijgt sterk* . Retrieved from: https://thebigstory.nl/social-media-in-nederland-in-2020-facebook-stijgt-licht-instagram-stijgt sterk/#:~:text=Gebruik%20naar%20leeftijd&text=Vooral%20Snapchat%2C%20Instagram%20en%20Tik,Youtube%20de%20meest%20gebruikte%20platforms.&text=De%20gemiddelde%20

Van den Berg, R. G. (2021). *SPSS One sample T-test.* Retrieved from: https://www.spss-tutorials.com/spss-one-sample-t-test/

Appendix 1 – Survey

Begin + demographic features

Dear participant,

Thank you for participating in this research. By participating you help me get closer to completing my bachelor's degree in Management, Society & Technology.

This research focuses on the influence of graphs appearances on the emotion of the reader. This research is specifically focused on the emotions and data visualization in the corona situation in the Netherlands. The findings of this research may help to find out what is the most pleasant way to depict situations with the help of graphs. Participating in this survey takes approximately 10 minutes.

Participation in this survey is voluntary and you can stop at any time. There is also always the option to indicate that you do not want to answer a question. When participating, no personal information about you will be stored. The only information that is stored and used is the answers you provide to the questions. It is in no way possible to link the answers to you personally.

If you have any questions or comments, you can contact me by sending an email to: m.ensink@student.utwente.nl

Thanks again!

Greetings, Merel Ensink

- o I have read and agreed to the text above

What is your gender?

- o Man
- o Woman
- o Else

What is your age?

What level of education do you follow? If you are not following an education currently, choose the highest education you have had.

- o Preparatory secondary vocational education
- o School of higher general secondary education
- o Pre-university education
- o Post-secondary vocational education
- o Higher vocational education
- o University (Bachelor)
- o University (Master)

Explanation of the research

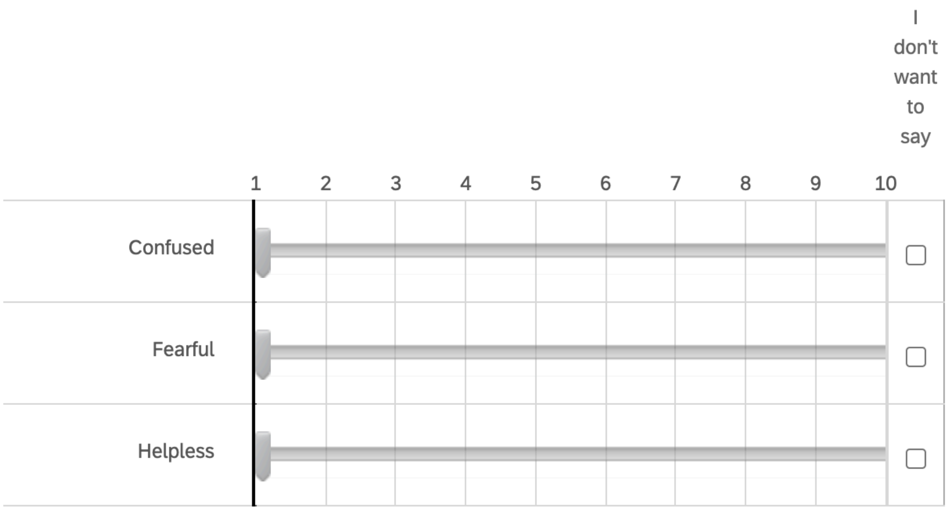
During this survey you will see graphs that describe the situation surrounding the coronavirus in the Netherlands. These situations can be, for example, the number of infections or the number of hospital admissions. After seeing the graphs, you should indicate whether and how much your emotions have been affected. There are 3 different emotions that are treated in this study, namely:

- **Confusion:** uncertainty on what is happening, what is intended and/or required.
- **Fear:** being afraid because of the threat of danger, pain or harm.
- **Helplessness:** feeling unable to help or change the situation.

First, you will be asked how much these emotions about the coronavirus in the Netherlands are present. After completing this question, you can click on the arrow at the bottom right to continue to the explanation of the following questions.

Rate on a scale of 1 - 10 how much these emotions are present in you right now. Here, 1 means that the emotion is not present at all and 10 that the emotion is very present.

You can answer this question by moving the button.



All the following questions have the same content, namely:

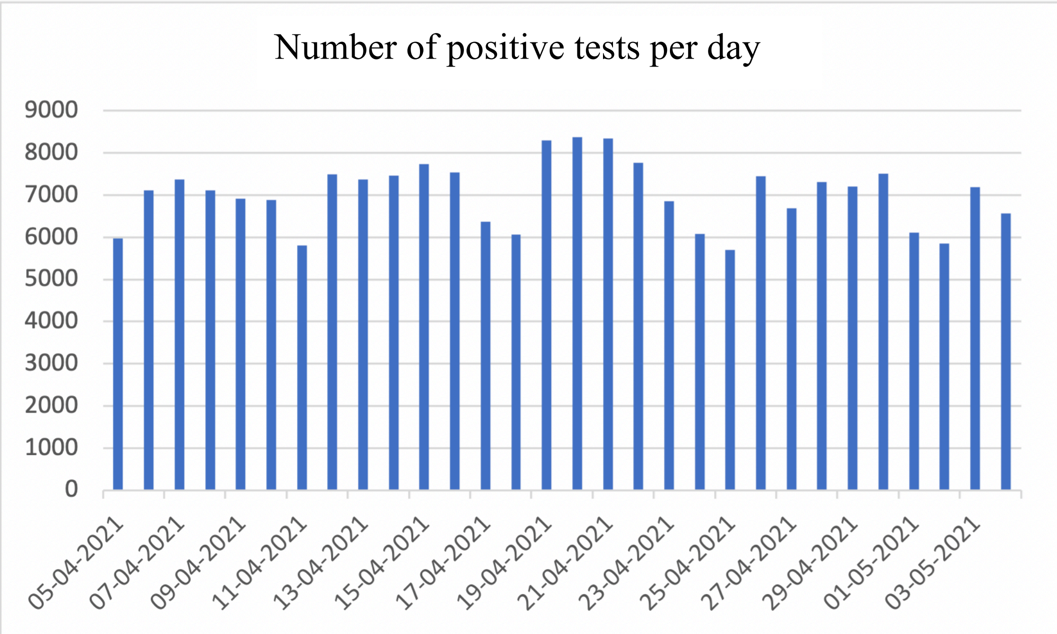
After seeing a graph, you must indicate on a scale of 1 - 5 whether the graph has influenced certain emotions. Here, 1 means that a emotion has been greatly reduced, 3 means that a emotion has remained the same and 5 means that a emotion has greatly increased. So, you compare how you felt about the coronavirus at the beginning of this survey with how you feel after seeing a graph.

Questions type of number/percentages (positive tests)

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



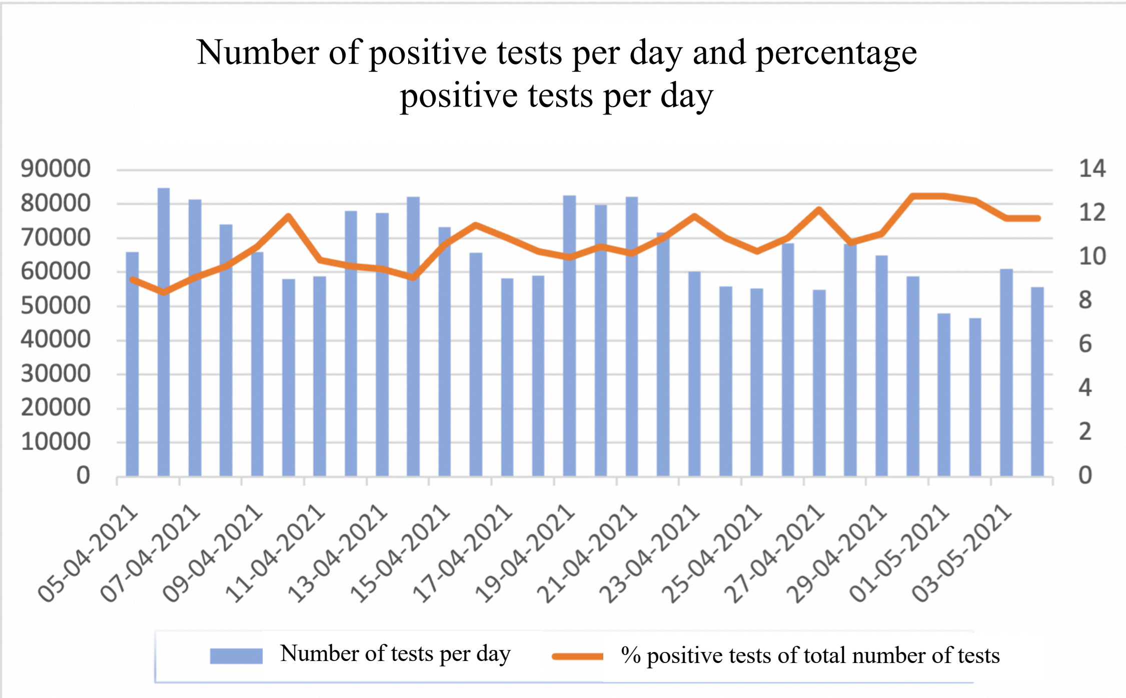
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



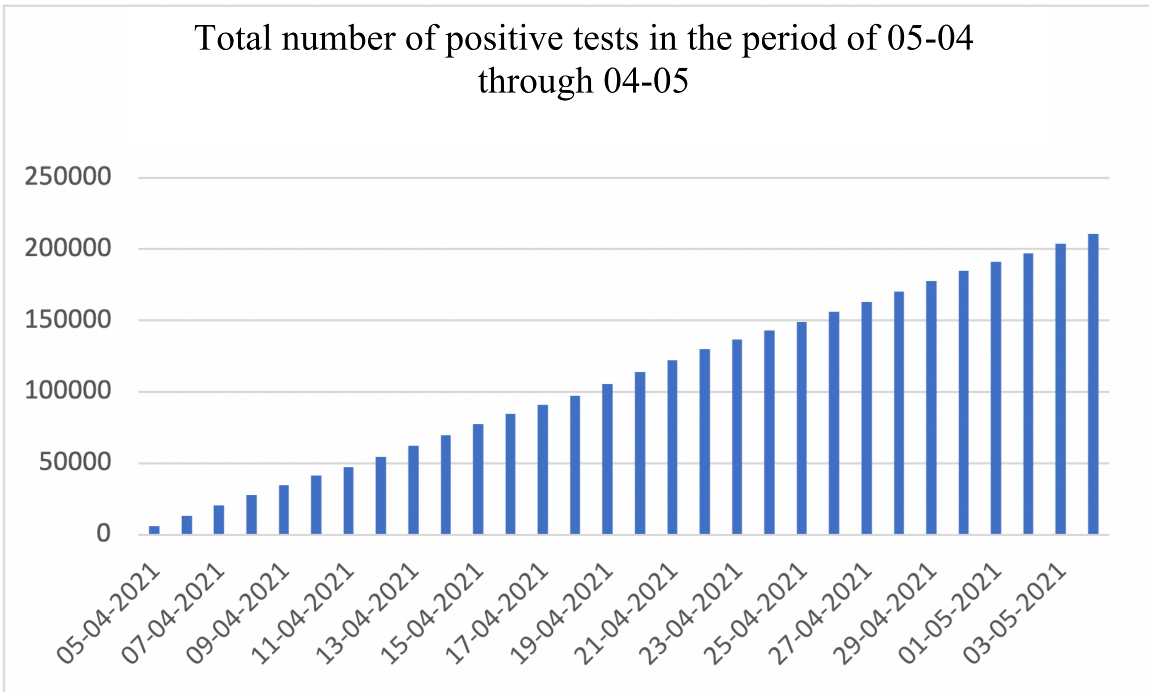
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



After seeing this graph, my feelings are:

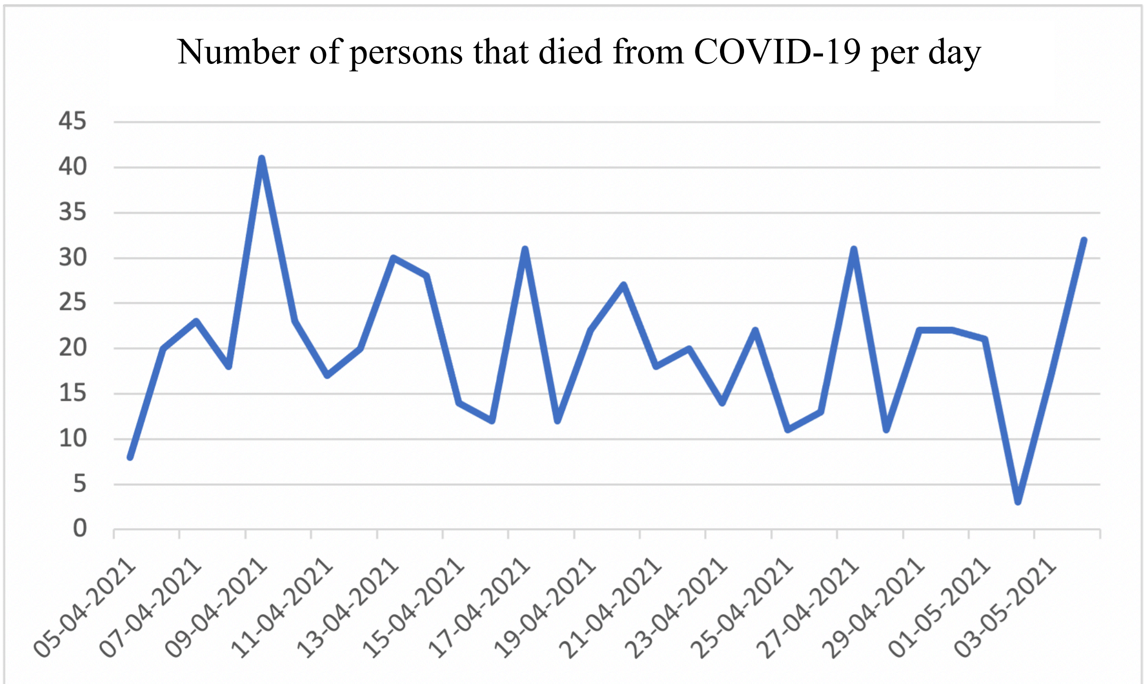
	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Questions scale (number of deceased persons) 1

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



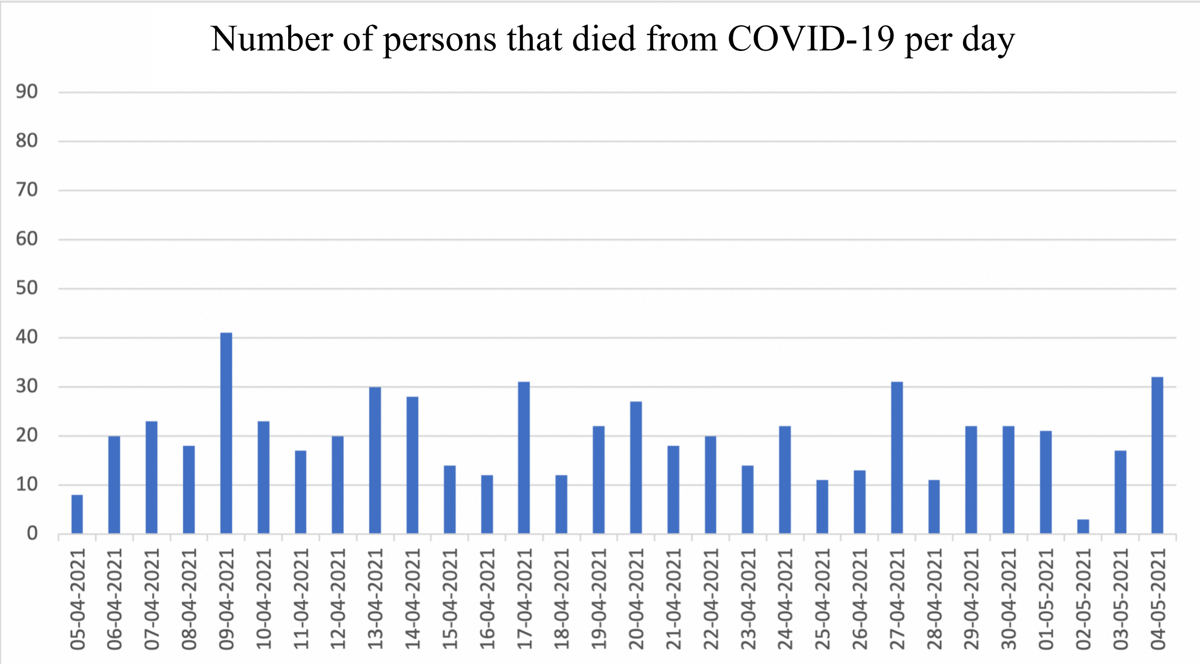
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



After seeing this graph, my feelings are:

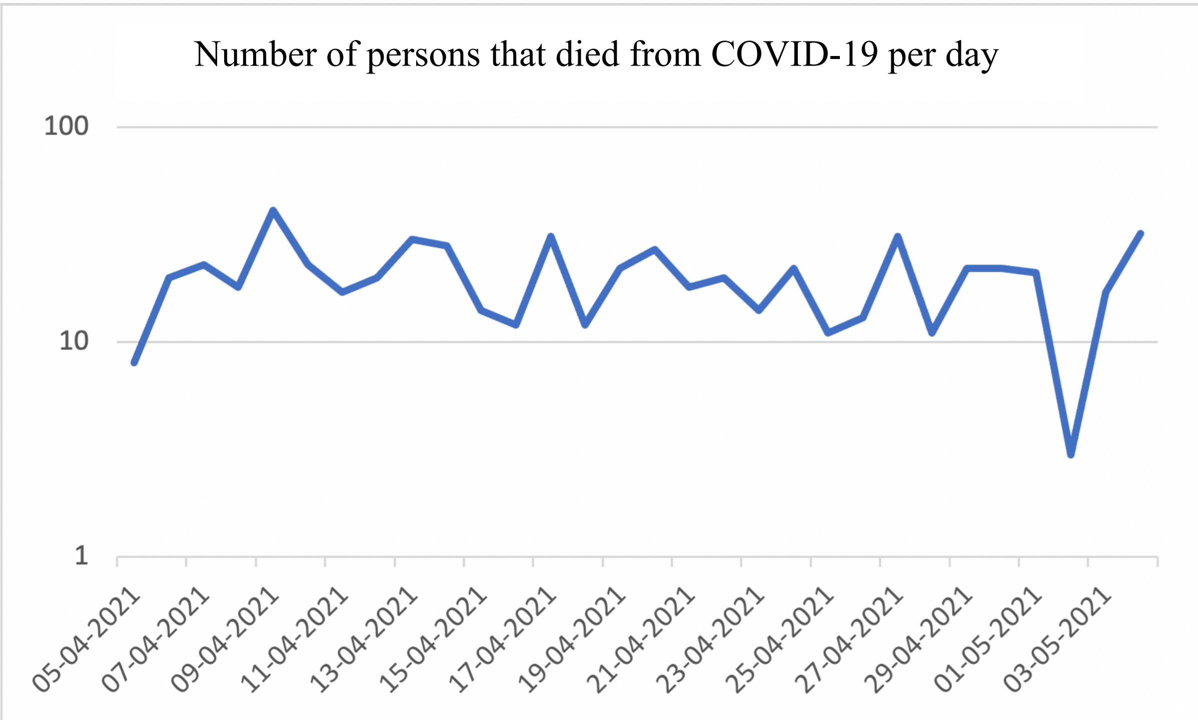
	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Questions scale (number of deceased persons) 2

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



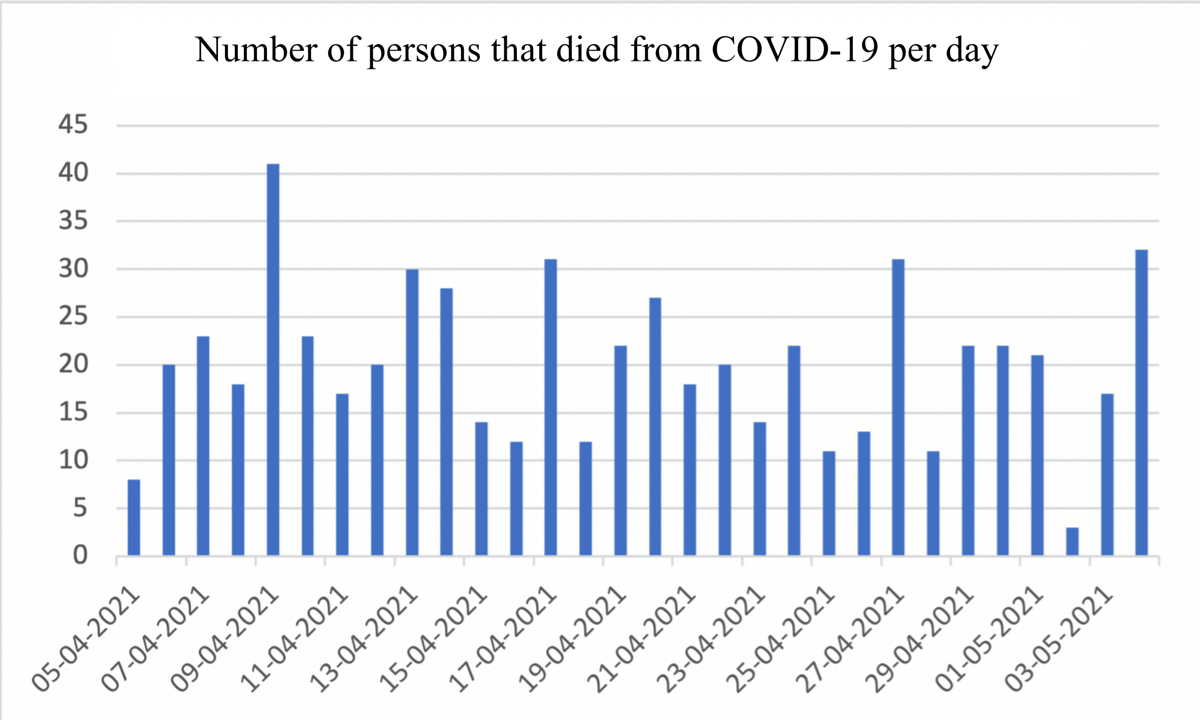
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



After seeing this graph, my feelings are:

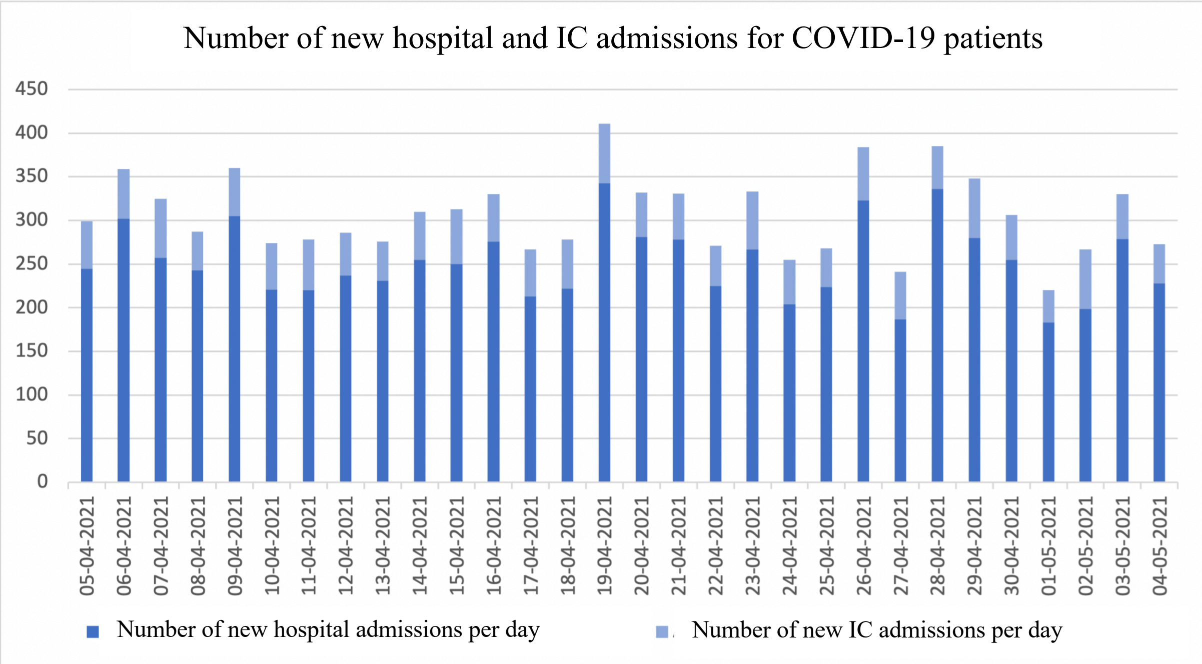
	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Questions color (number of hospital admissions) 1

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



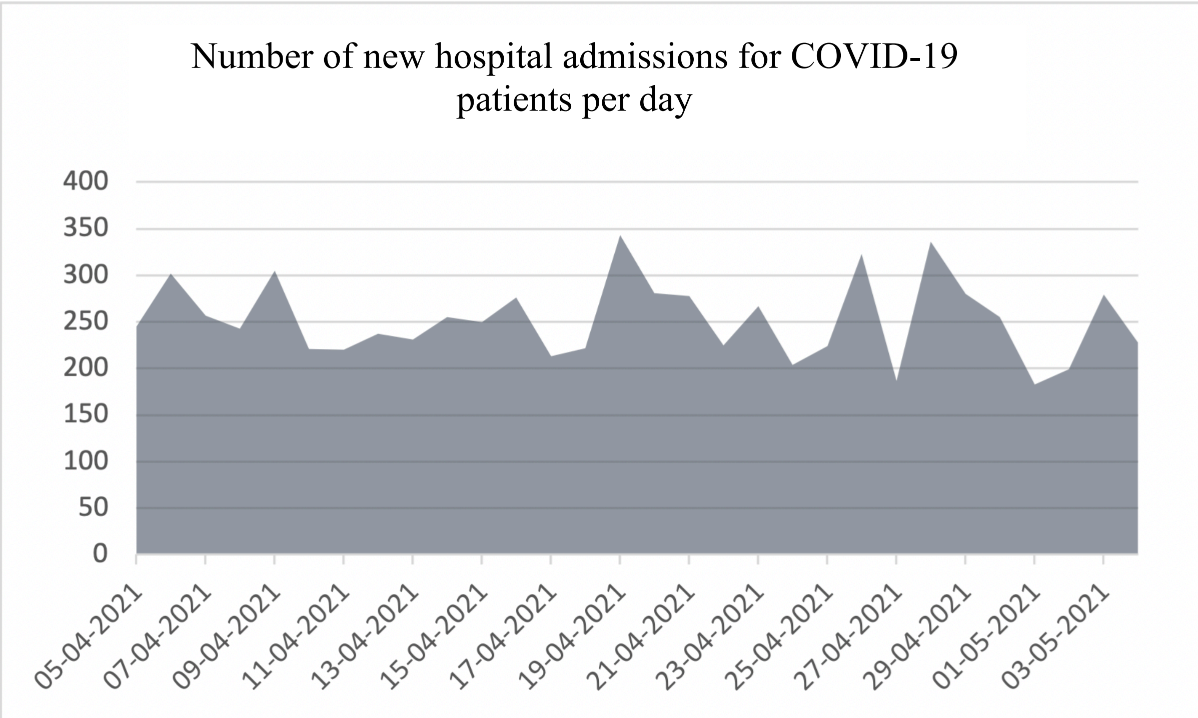
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



After seeing this graph, my feelings are:

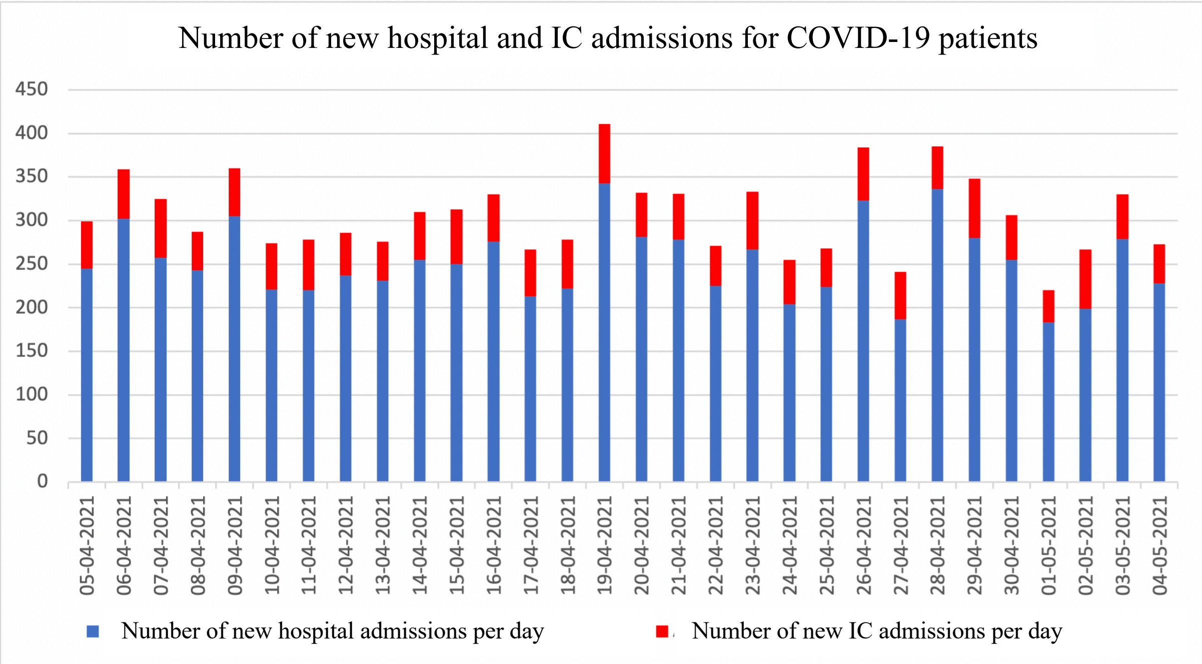
	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Questions color (number of hospital admissions) 2

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



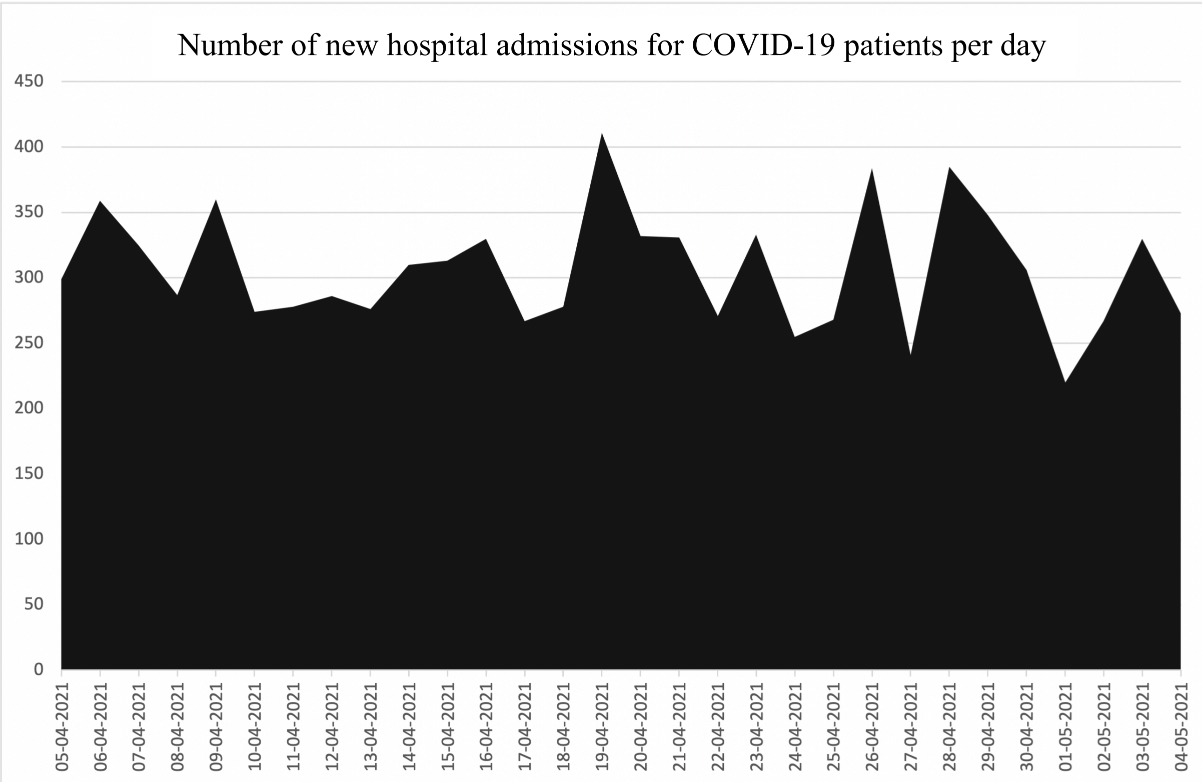
After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Note: the question is about the comparison between your answer initially and after seeing the graph. Your first answer was*:

Confused: *Displays answer giving at first question*
Fear: *Displays answer giving at first question*
Helpless: *Displays answer giving at first question*

*On a scale of 1 – 10, where 1 means that the feeling was not present at all and 10 means that a feeling was greatly present.



After seeing this graph, my feelings are:

	1. Greatly reduced	2. Reduced	3. Unchanged	4. Increased	5. Greatly increased	6. I don't want to tell
Confusion	○	○	○	○	○	○
Fear	○	○	○	○	○	○
Helplessness	○	○	○	○	○	○

Appendix 2 – Graphs in the corona report

Visualization of corona data	Color	Scale	Type of numbers/ percentages	Page number
Number of persons with a positive test per day	**Old**: purple **New**: yellow	**Min:** 0 **Max:** 7.500 **Interval:** 2.500	Absolute numbers	6
Number of deceased persons due to corona per day	**Old**: purple **New**: yellow	**Min:** 0 **Max:** 60 **Interval:** 20	Absolute numbers	6
Number of new hospital admissions per day	Purple, new data is separated by a dotted line	**Min:** 0 **Max:** 300 **Interval:** 100	Absolute numbers	8
Number of new Intensive Care admissions per day	Purple, new data is separated by a dotted line	**Min:** 0 **Max:** 60 **Interval:** 20	Absolute numbers	8
Number of positive tests per age category (9 graphs)	**Old**: purple **New**: yellow	**Min:** 0 **Max:** 1.500 **Interval:** 500	Absolute numbers	9
Number of positive tests per safety region (25 graphs)	**Old**: purple **New**: yellow	**Min:** 0 **Max:** 750 **Interval:** 250	Absolute numbers	10 – 12
Number of deceased persons per age category (9 graphs)	**Old**: light purple **New**: light yellow	**Min:** 0 **Max:** 50 **Interval:** 10	Absolute numbers	13
Number of deceased persons per safety region (25 graphs)	**Old**: light purple **New**: light yellow	**Min:** 0 **Max:** 8 **Interval:** 2	Absolute numbers	14 – 16
Age and male-female distribution of the number of positive tests	**Man**: blue **Women**: dark pink	**Min:** 0 **Max:** 40.000 **Interval:** 10.000 **x-axis:** 4 years per age category	Absolute numbers	22
Age and male-female distribution of the number of deceased persons	**Man**: light blue **Women**: light pink **Not mentioned**: light gray	**Min:** 0 **Max:** 600 **Interval:** 200 **x-axis:** 4 years per age category	Absolute numbers	22
Number of positive tests in nursing homes and residential care centers	**Old:** purple **New:** Yellow	**Min:** 0 **Max:** 200 **Interval:** 50	Absolute numbers	30

Number of deceased persons in nursing homes and residential care centers	**Old:** light purple **New:** Light yellow	**Min:** 0 **Max:** 40 **Interval:** 10	Absolute numbers	30
Number of nursing homes and residential care centers with at least one infection	**Old:** purple **New:** yellow	**Min:** 0 **Max:** 15 **Interval:** 5	Absolute numbers	31
Number of positive tests in disabled care institutions	**Old:** purple **New:** yellow	**Min:** 0 **Max:** 80 **Interval:** 20	Absolute numbers	32
Number of deceased persons in disabled care institutions	**Old:** light purple **New:** light yellow	**Min:** 0 **Max:** 3 **Interval:** 1	Absolute numbers	32
Number of disabled care institutions with at least one infection	**Old:** purple **New:** yellow	**Min:** 0 **Max:** 20 **Interval**: 5	Absolute numbers	33
Number of positive tests for independently living people of 70 years or older	**Old:** purple **New:** yellow	**Min:** 0 **Max:** 500 **Interval:** 100	Absolute numbers	34
Number of deceased persons of 70 years or older independently living	**Old:** light purple **New:** light yellow	**Min:** 0 **Max:** 25 **Interval:** 5	Absolute numbers	34
Number of tests and percentage positive tests per week	**Number of tests:** gray **Percentage positive tests:** purple	**Min:** 0 **Max:** 400.000 **Interval:** 200.000 **Min:** 0 **Max:** 30 **Interval:** 5	Absolute numbers Percentages	36
Number of tests and percentage positive tests per week and age category (9 graphs)	**Number of tests:** gray **Percentage positive tests:** purple	**Min:** 0 **Max:** 6.000 **Interval:** 2.000 **Min:** 0 **Max:** 30 **Interval:** 5	Absolute numbers Percentages	39
Number of tests and percentage positive tests per week and target audience (9 graphs)	**Number of tests:** gray **Percentage positive tests:** purple	**Min:** 0 **Max:** 90.000 **Interval:** 30.000 **Min:** 0 **Max:** 30 **Interval:** 5	Absolute numbers Percentages	42

Number of tests and percentage positive tests per week and target audience (9 graphs)	**Number of tests:** gray **Percentage positive tests:** purple	**Min:** 0 **Max:** 90.000 **Interval:** 30.000 **Min:** 0 **Max:** 30 **Interval:** 5	Absolute numbers Percentages	43
Number of tests and percentage positive tests per week and safety region (25 graphs)	**Number of tests:** gray **Percentage positive tests:** purple	**Min:** 0 **Max:** 3.000 **Interval:** 1.000 **Min:** 0 **Max:** 30 **Interval:** 5	Absolute numbers Percentages	44 – 45
Estimated number of contagious persons	**Most likely number:** dark purple **95% confident interval:** light purple	**Min:** 0 **Max:** 200.000 **Interval:** 50.000	Absolute numbers	48
Number of infected persons to report date, with expected extra cases.	**Infected in the Netherlands:** Blue **Infected outside the Netherlands:** Yellow **Expected extra cases:** Light gray **First day of illness reports:** Green	**Min:** 0 **Max:** 12.000 **Interval:** 4.000	Absolute numbers	49
The effective reproduction number R	**Most likely number:** dark purple **95% confident interval:** light purple	**Min:** 0 **Max:** 3 **Interval**: 1	Reproduction number	49
Percentage of positive tested persons per week, based on virology laboratories	Dark pink	**Min:** 0 **Max:** 40 **Interval:** 10	Percentages	51

Number of patients with flu-like symptoms tested positive for one of the viruses or negative	**Negative:** turquoise **Influenza:** Purple **RSV:** Blue **Rhinovirus:** Gray **Enterovirus:** yellow **COVID-19:** Dark pink	**Min:** 0 **Max:** 100 **Interval:** 50	Absolute numbers	55
Percentage of patients with flu-like symptoms tested positive for one of the viruses (5 graphs)	**Influenza:** Purple **RSV:** Blue **Rhinovirus:** Gray **Enterovirus:** Yellow **COVID-19:** Dark pink	**Min:** 0 **Max:** 40 **Interval:** 10 Dots indicate the number per week, a line indicates its 5-week average.	Percentages	56
Percentage of people with COVID-19 like symptoms from weekly survey	Purple	**Min:** 0 **Max:** 4 **Interval:** 1 Dots indicate the percentage per day, a line indicates its 5-day average	Percentages	57
Total mortality per week compared to the number expected on the basis of previous years	**Total:** Purple **Expected:** Light yellow	**Min:** 2.000 **Max:** 5.000 **Interval:** 1.000	Absolute numbers	58

Table 9: graphs in corona report (27th of April)

Appendix 3 – Graphs on coronavirus dashboard

Visualization of corona data	**Color**	**Scale**	**Type of numbers/ percentages**	**Heading**
Total number of vaccinations	Light blue	**Min:** 0 **Max:** 20.000.000 **Interval:** 10.000.000	Cumulative numbers	Vaccinations
Hospital admissions over time, by date of admission (total)	**Admissions**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 800 **Interval:** 200	Absolute numbers	Hospitals
Hospital admissions over time, by date of admission (last 5 weeks)	**Admissions**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 250 **Interval:** 50	Absolute numbers	Hospitals
Hospital admissions over time per age category (total)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 1.000 **Interval:** 200	Relative numbers	Hospitals
Hospital admissions over time per age category (last 5 weeks)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 250 **Interval:** 50	Relative numbers	Hospitals

Occupancy of ordinary hospital beds (without IC) over time (total)	**Occupancy beds:** light blue **Uncertain:** gray	**Min:** 0 **Max:** 4.000 **Interval:** 1.000	Absolute numbers	Hospitals
Occupancy of ordinary hospital beds (without IC) over time (last 5 weeks)	**Occupancy beds:** light blue	**Min:** 0 **Max:** 2.000 **Interval:** 500	Absolute numbers	Hospitals
Admissions intensive care over time (total)	**Admissions**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 150 **Interval:** 50	Absolute numbers	Hospitals
Admissions intensive care over time (last 5 weeks)	**Admissions**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 50 **Interval:** 10	Absolute numbers	Hospitals
Intensive care admissions over time per age category (total)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 200 **Interval:** 50	Relative numbers	Hospitals
Hospital admissions over time per age category (last 5 weeks)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 60 **Interval:** 20	Relative numbers	Hospitals
Occupancy of intensive care beds over time (total)	**Occupancy beds:** light blue **Uncertain:** gray	**Min:** 0 **Max:** 1.500 **Interval:** 500	Absolute numbers	Hospitals

Occupancy of intensive care beds over time (last 5 weeks)	**Occupancy beds:** light blue	**Min:** 0 **Max:** 1.000 **Interval:** 200	Absolute numbers	Hospitals
Number of positive tested persons over time (per 100.000 inhabitants) (total)	**Positive tests:** light blue **Positive tests average over 7 days:** dark blue	**Min:** 0 **Max:** 80 **Interval:** 20	Relative numbers	Infections
Number of positive tested persons over time (per 100.000 inhabitants) (last 5 weeks)	**Positive tests:** light blue **Positive tests average over 7 days:** dark blue	**Min:** 0 **Max:** 50 **Interval:** 10	Relative numbers	Infections
Positive tests over time per age category (total)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 150 **Interval:** 50	Relative numbers	Infections
Positive tests over time per age category (last 5 weeks)	**0 -19:** Blue **20-29:** Turquoise **30-39:** Green **40-49:** Yellow **50-59:** Ochreous **60-69:** Orange **70-79:** Brown **80-89:** Pink **90+:** Purple **Total:** dotted line **Uncertain:** gray	**Min:** 0 **Max:** 80 **Interval:** 20	Relative numbers	Infections
Development of positive tests (growth rate)	Blue	**Min:** 0 **Max:** -40 **Interval:** -10	Relative numbers	Infections

Percentage positive tests over time (total)	**Percentage positive tests:** light blue **Average percentage positive tests over 7 days:** dark blue	**Min:** 0 **Max:** 25 **Interval:** 5	Relative numbers	Infections
Percentage positive tests over time (last 5 weeks)	**Percentage positive tests:** light blue **Average percentage positive tests over 7 days:** dark blue	**Min:** 0 **Max:** 15 **Interval:** 5	Relative numbers	Infections
Number of tests over time (total)	**Tests with a result:** gray **Average number of tests with a result over 7 days:** dark blue **Tests with a positive result:** light blue **Average number of tests with a positive result over 7 days:** blue	**Min:** 0 **Max:** 100.000 **Interval:** 20.000	Absolute numbers	Infections
Number of tests over time (last 5 weeks)	**Tests with a result:** gray **Average number of tests with a result over 7 days:** dark blue **Tests with a positive result:** light blue **Average number of tests with a positive result over 7 days:** blue	**Min:** 0 **Max:** 60.000 **Interval:** 20.000	Absolute numbers	Infections

Number of contagious people over time (total)	**Estimated number of contagious people:** dark blue **Margin of uncertainty:** light blue	**Min:** 0 **Max:** 250.000 **Interval:** 50.000	Absolute numbers	Infections
Number of contagious people over time (last 5 weeks)	**Estimated number of contagious people:** dark blue **Margin of uncertainty:** light blue	**Min:** 0 **Max:** 200.000 **Interval:** 50.000	Absolute numbers	Infections
Reproduction number (total)	Blue	**Min:** 0 **Max:** 2,5 **Interval:** 0,5	Reproduction number	Infections
Reproduction number (last 5 weeks)	Blue	**Min:** 0 **Max:** 2 **Interval:** 0,5	Reproduction number	Infections
Patient who died from COVID-19 over time (total)	**Deceased persons**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 250 **Interval:** 50	Absolute numbers	Infections
Patient who died from COVID-19 over time (last 5 weeks)	**Deceased persons**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 50 **Interval:** 10	Absolute numbers	Infections
COVID-19 mortality: breakdown by age	**Distribution of age category:** gray **Deceased persons per age category:** blue	**Min:** 0 **Max:** 40 **Interval:** 20	Relative numbers	Infections

Number of deceased persons per week	**Expected number:** blue **Actual number:** dark blue **Upper and lower limits of expected number:** light blue	**Min:** 0 **Max:** 6.000 **Interval:** 2.000	Absolute numbers	Infections
Codes of conduct over time (following or supporting)	**Following:** blue **Supporting:** yellow	**Min:** 0 **Max:** 100 **Interval:** 25	Relative numbers	Behavior
Number of positively tested nursing home residents over time (total)	**Number of positive tested residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 500 **Interval:** 100	Absolute numbers	Vulnerable groups
Number of positively tested nursing home residents over time (last 5 weeks)	**Number of positive tested residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 25 **Interval:** 5	Absolute numbers	Vulnerable groups
Number of infected nursing homes over time (total)	Blue	**Min:** 0 **Max:** 1.000 **Interval:** 200	Absolute numbers	Vulnerable groups
Number of infected nursing homes over time (last 5 weeks)	Blue	**Min:** 0 **Max:** 200 **Interval:** 50	Absolute numbers	Vulnerable groups
Number of deceased nursing home residents over time (total)	**Number of deceased residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 100 **Interval:** 20	Absolute numbers	Vulnerable groups

Number of deceased nursing home residents over time (last 5 weeks)	**Number of deceased residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 6 **Interval:** 2	Absolute numbers	Vulnerable groups
Number of positively tested residents of disabled care institutions (total)	**Number of positive tested residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 120 **Interval:** 20	Absolute numbers	Vulnerable groups
Number of positively tested residents of disabled care institutions (last 5 weeks)	**Number of positive tested residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 10 **Interval:** 2	Absolute numbers	Vulnerable groups
Total number of infected disabled care institutions (total)	Blue	**Min:** 0 **Max:** 250 **Interval:** 50	Absolute numbers	Vulnerable groups
Total number of infected disabled care institutions (last 5 weeks)	Blue	**Min:** 0 **Max:** 50 **Interval:** 10	Absolute numbers	Vulnerable groups
Number of deceased disabled care institutions residents (total)	**Number of deceased residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 4 **Interval:** 1	Absolute numbers	Vulnerable groups

Number of deceased disabled care institutions residents (total)	**Number of deceased residents**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 1 **Interval:** 0.2	Absolute numbers	Vulnerable groups
New cases among persons older than 70 that live at home (total)	**Number of positive tests**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 1.200 **Interval:** 200	Absolute numbers	Vulnerable groups
New cases among persons older than 70 that live at home (last 5 weeks)	**Number of positive tests**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 400 **Interval:** 100	Absolute numbers	Vulnerable groups
Number of deceased persons that are 70 years or older and live at home (total)	**Number of deceased persons**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 50 **Interval:** 10	Absolute numbers	Vulnerable groups
Number of deceased persons that are 70 years or older and live at home (last 5 weeks)	**Number of deceased persons**: light blue **Average over 7 days**: dark blue **Uncertain:** gray	**Min:** 0 **Max:** 15 **Interval:** 5	Absolute numbers	Vulnerable groups
Average number of virus particles over time in wastewater (total)	Blue	**Min:** 0 **Max:** 800 **Interval:** 200 x 100 billion	Relative numbers	Early signals

Average number of virus particles over time in wastewater (last 5 weeks)	Blue	**Min:** 0 **Max:** 300 **Interval:** 100 x 100 billion	Relative numbers	Early signals
Number of patients with symptoms that could indicate COVID-19 per 100.000 inhabitants (total)	Blue	**Min:** 0 **Max:** 120 **Interval:** 20	Relative numbers	Early signals
Number of patients with symptoms that could indicate COVID-19 per 100.000 inhabitants (last 5 weeks)	Blue	**Min:** 0 **Max:** 10 **Interval:** 2	Relative numbers	Early signals
Warnings from the corona detector app (total)	Blue	**Min:** 0 **Max:** 2.000 **Interval:** 500	Absolute number	Other
Warnings from the corona detector app (last 5 weeks)	Blue	**Min:** 0 **Max:** 800 **Interval:** 200	Absolute number	Other

Table 10: graphs on coronavirus dashboard (14th of June)

Appendix 4 – Figures on subpage

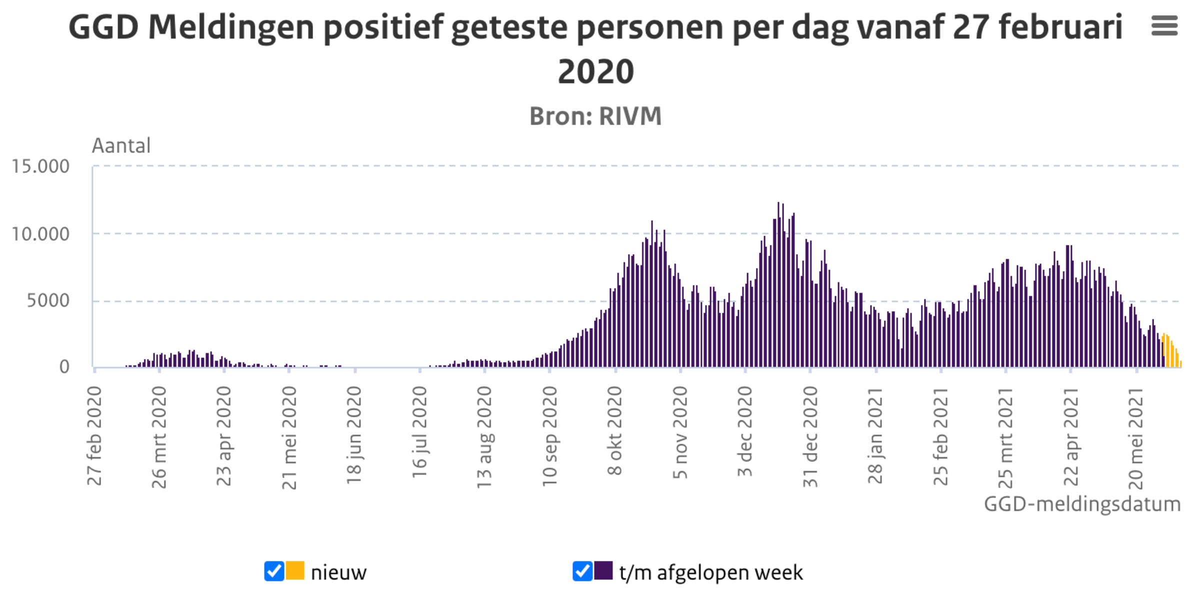


Figure 9: Positively tested persons per day

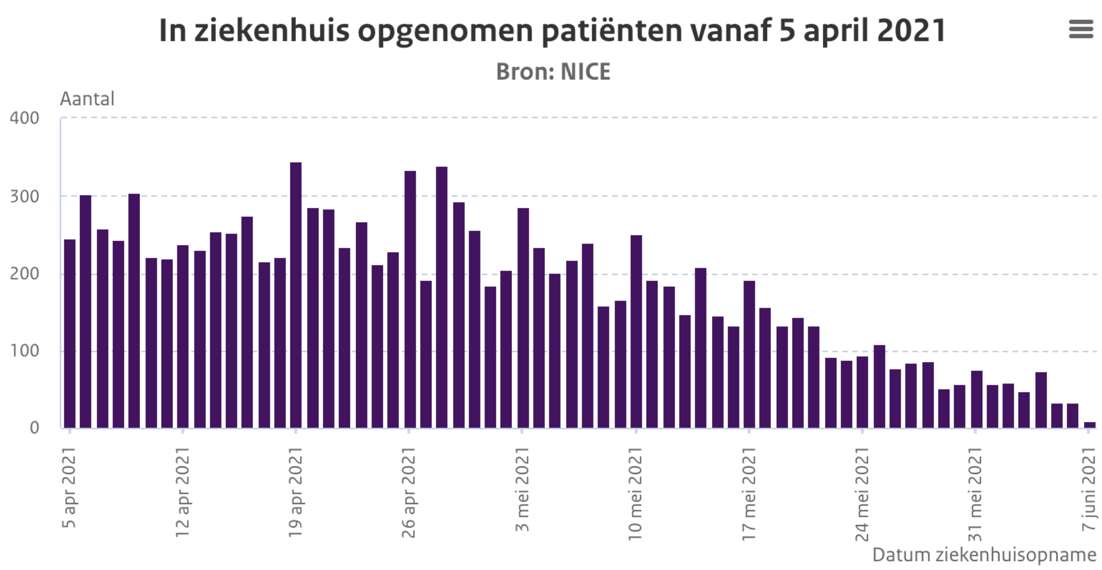


Figure 10: Number of hospital admissions per day.

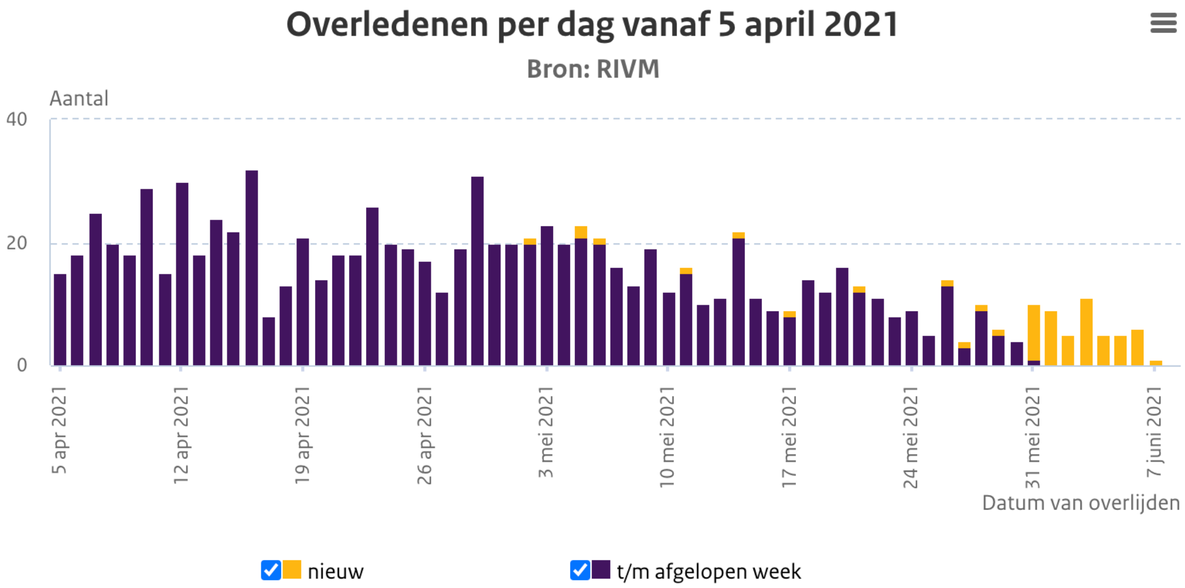


Figure 11: Number of deceased persons per day

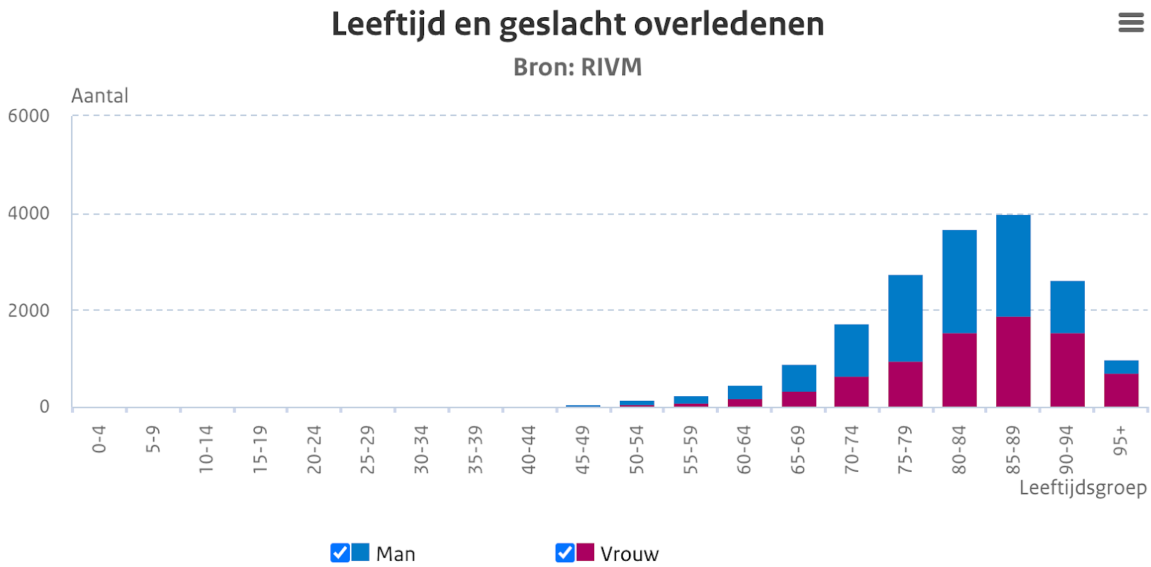


Figure 12: Number of deceased persons per age and gender

Appendix 5 - Calculation gender

100 women: 99 men

Since the CBS does not use the category “else”, these are left out of the sample for calculation. Gives:

CBS distribution:

100 women and 99 men = 199 total

Meaning that for a total of 98:
Women: 100/199*98 » 49,25
Men: 99/199*98 » 48,75

Appendix 6 – Distribution age

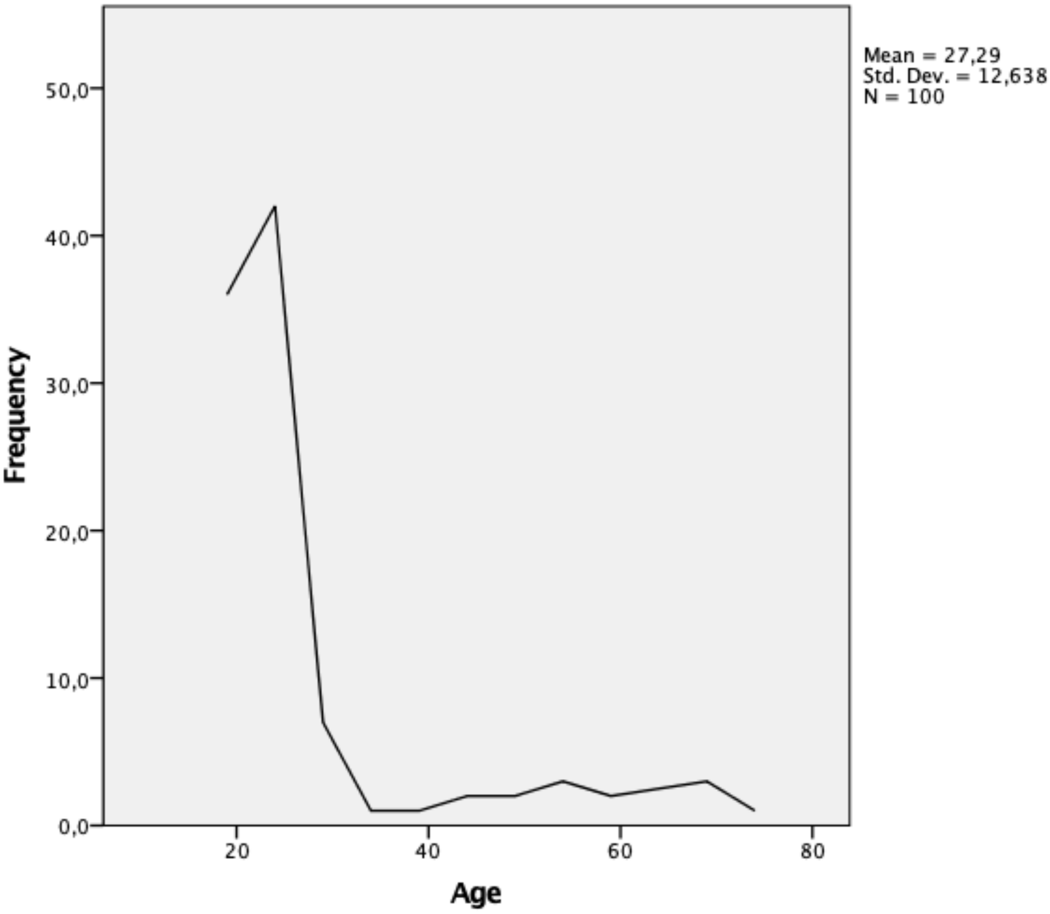


Figure 13: distribution of age in the response

Appendix 7 – Descriptives for the variables

Color

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	50	2	5	3,20	,670	1,018	,337
Fear	50	2	4	3,02	,428	,125	,337
Helplessness	50	1	4	3,10	,544	-,713	,337
Valid N (listwise)	50						

Table 11: descriptives for the graph with similar colors

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	50	2	5	3,16	,584	,615	,337
Fear	50	2	4	3,06	,470	,213	,337
Helplessness	50	2	4	3,12	,480	,357	,337
Valid N (listwise)	50						

Table 12: descriptives for the graph with a light color

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	50	1	5	3,26	,777	,046	,337
Fear	50	1	4	3,00	,606	-,573	,337
Helplessness	50	1	4	2,94	,550	-,808	,337
Valid N (listwise)	50						

Table 13: Descriptives of the graph with different colors

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	50	1	5	3,50	,886	,092	,337
Fear	50	1	5	3,32	,794	,115	,337
Helplessness	50	1	5	3,14	,639	,363	,337
Valid N (listwise)	50						

Table 14: Descriptives for the graph with dark colors

Type of scale

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	52	2	5	3,19	,658	,637	,330
Fear	52	2	4	3,17	,513	,264	,330
Helplessness	52	2	4	3,29	,498	,458	,330
Valid N (listwise)	52						

Table 15: Descriptives for the graph with a linear scale

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	52	1	4	2,90	,495	-1,231	,330
Fear	52	2	4	3,00	,524	,000	,330
Helplessness	52	2	4	3,10	,454	,418	,330
Valid N (listwise)	52						

Table 16: Descriptives for the graph with a big scale

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	48	1	5	3,00	,715	,000	,343
Fear	48	1	4	3,00	,684	-,416	,343
Helplessness	48	1	5	3,17	,663	-,195	,343
Valid N (listwise)	48						

Table 17.: Descriptives for the graph with a logarithmic scale

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	48	1	5	2,98	,699	-,361	,343
Fear	48	2	4	3,19	,532	,176	,343
Helplessness	48	1	4	3,13	,531	-,751	,343
Valid N (listwise)	48						

Table 18: descriptives for the graph with a normal scale

Type of number

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	100	1	4	3,00	,492	-1,036	,241
Fear	100	1	4	3,03	,521	-,831	,241
Helplessness	100	1	5	3,21	,608	-,137	,241
Valid N (listwise)	100						

Table 19: descriptives for the graph with absolute numbers

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	100	1	5	3,20	,816	,068	,241
Fear	100	1	5	3,00	,682	-,586	,241
Helplessness	100	1	5	3,03	,611	-,286	,241
Valid N (listwise)	100						

Table 20: descriptives for the graph with relative numbers

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
Confusion	100	2	5	3,34	,699	-,221	,241
Fear	100	2	5	3,38	,616	,339	,241
Helplessness	100	2	5	3,47	,594	,561	,241
Valid N (listwise)	100						

Table 21: descriptives for the graph with cumulative numbers