Universiteit Leiden

Educating vocational education students on misleading data visualisations.

Maaike Boele (2340887)

SCS Research Report, Biology, Leiden University, Supervisor: Dr.W.Wijnker, Project duration 6-2-2023/12-5-2023 (16 EC + 3 EC research proposal), contact person raw data: Dr.W.Wijnker

Abstract

Data visualisations are often used to communicate to the general public about complex issues. Although they are supposed to objectively portray data, media sometimes use graphs to frame a specific story which can purposefully or accidentally lead to misinterpretation of the data by their viewers. Not everyone is able to recognize misleading graphs, especially students of vocational education programmes struggle with this. To prevent misinformation from spreading in this group, it is important that effective education strategies are developed. To investigate how to increase graph literacy of vocational education students, a lesson on misleading data visualisations was developed based on existing knowledge of educational strategies. The extent to which the lesson contributed to students' knowledge on misleading graphs was investigated using two surveys; one at the start and one at the end of the lesson. Furthermore, a teacher manual was developed to aid teachers to continue educating students on misleading data visualisations. Interviews with teachers were performed to study how the manual can help a teacher to continue education on misleading graphs. Results indicated that a short lesson on misleading data visualisations can, to some extent, improve students' ability to recognize misleading graphs. A teacher manual can support teachers with preparing lessons on misleading graphs to further increase students' graph literacy.

Keywords: graph literacy, data visualization, graph education, teacher support

Introduction

Data visualisations are increasingly abundant in our digitizing world. Graphs make data more accessible to the general public (B. Lee et al., 2020) and are used by scientists and the mass media to communicate about complex problems or change people's attitude towards a subject (Pandey et al., 2014). They allow for direct comparison of measurements, show trends over time, and highlight correlations in data (Shah & Hoeffner, 2002). We saw clear examples of the use of data visualisations to communicate to the general public during the COVID pandemic, where data visualisations were widely used by the government and scientists to portray the importance of lockdown measures (Jayasinghe et al., 2020; Romano et al., 2020). However, there is a serious drawback to using data visualisations: figures can by accident or on purpose present information incorrectly.

Misleading data visualisations are not new to us. Already in 1950, misleading graphs were often used in newspapers (Huff, 1954). Because of the internet, personal computers, and social media we now have access to a much larger data source. People encounter graphs in magazines, on their energy bills, and in Twitter feeds. This data overload makes it hard to distinguish between correct and incorrect data visualisations. It is often up to users to judge if a data source can be trusted. Their ability to recognize misleading graphs is dependent on multiple factors such as graph design and perception.

Not all data visualisations are well-designed, or easy to interpret. Common mistakes that intentionally or accidentally misinform readers, are using an incorrect graph type to visualise data or cherry-picking data (i.e. leaving out inconvenient trends) (Nguyen et al., 2021). Furthermore, design choices within a graph can also lead to misinterpretation of data. Examples of this are truncated y-axes (i.e. omitted baseline) (Yang et al., 2021) and confusing colour choices (e.g. going against colour norm).

On top of problems with the design of data visualisations, not everyone is able to correctly interpret visually represented data. According to the OECD (2013), younger age groups, especially those with a vocational education, have more difficulty correctly interpreting data visualisations, compared to adults. To prevent these groups from being misinformed by misleading graphs, educational programmes should spend more time on teaching about data visualisations. Different educational strategies (e.g. inquiry-based learning, activating prior knowledge, and scaffolding) have been proven to be useful when educating vocational education students. However, little is known about the effectiveness of these teaching strategies when using them to educate students on misleading data visualisations (Börner et al., 2019). To that end, I designed a short introductory lesson on data visualisations and investigated to what extent the developed lesson can increase graph literacy of vocational education students.

In the theoretical framework, I discuss five important aspects to consider when designing educational programmes aimed at increasing students' graph literacy. First, I look at the current graph literacy of (young) adults in the Netherlands to understand the need for education on data visualisations. Second, I discuss the definition of graph literacy and concentrate on methods to measure graph literacy. Third, I focus on graph characteristics to understand how correct graphs are designed. Fourth, I concentrate on the misuse of graph characteristics, in relation to graph literacy skills, to understand how misinterpretations of graphs can arise. Last, I describe educational strategies that can be used to design a lesson on data visualisations to avoid misinterpretations.

Theoretical framework

Numeracy versus graph literacy

As people regularly encounter data visualisations, in a personal and professional settings, that tell them important information needed to perform certain tasks, it is important to enable them to correctly interpret data visualisations. Although numeracy and graph literacy are correlated (Nayak et al., 2016), to correctly interpret graphs, people not only need sufficient numeracy skills, but they also need specific skills to interpret data visualisations. It was found that higher numeracy has a negative correlation with susceptibility to misinformation by misleading data visualisations (van der Linden, 2022; Nayak et al., 2016) This suggests that to limit misinterpretations of misleading data visualisations, it is crucial to specifically educate groups with a low numeracy on misleading data visualisations.

Details on graph literacy of the Dutch population are lacking. Multiple studies did investigate the numeracy of adults in the Netherlands. Compared to other countries, adults in the Netherlands have an above-average numeracy (i.e. understanding of the real number line, time, ratio concepts, fractions, percentages, probabilities, measurement, and estimation (Reyna et al., 2009)). However, within the Netherlands, there is a significant difference in numeracy between different age groups and across education. According to the OECD (2013), different minorities in the Dutch population have poor numeracy, with 13.2% of adults scoring below average in numeracy. On top of that, the age group 16-24 has a lower numeracy compared to adults between 25-44 (OECD, 2013). Specifically, graduates of vocational education programmes have a lower proficiency in numeracy compared to young adults of the same age who graduated from secondary educational programmes (Galesic & Garcia-Retamero, 2011b; Research centrum voor onderwijs en arbeidsmarkt, 2019).

Measuring graph literacy

Graph literacy often is referred to as “the ability and skill of people to read and interpret visually represented data to extract information from data visualisation” (S. Lee et al., 2017, pp 532). This is often measured using the Short Graph Literacy (SGL) scale (Okan et al., 2019) or the Objective Graph Literacy (OGL) scale (Galesic & Garcia-Retamero, 2011a). However, correctly interpreting graphical representations of data is more complex in the case of misleading data visualisations, often leading to wrong conclusions. Therefore, I suggest including “the ability to recognize and explain mistakes in data visualisations” to above-described definition. By teaching students the most common mistakes, it is expected that it becomes easier for students to recognize misleading data visualisations in the media, helping with correctly interpreting those visualisations. To that end, for this study, I expand above-described definition of graph literacy to the ability and skill to read and interpret visually represented data, “to extract information from data visualisations, and the ability to recognize common mistakes in graph design and explain why certain mistakes can cause misinterpretations”.

To measure the effect of a lesson on graph literacy following the definition of this study using SGL or OGL would not suffice, as this does not allow to measure the ability of students to recognize and explain common mistakes in graphs. Students’ ability to notice a mistake in visualisations can be tested by showing them multiple graphs and letting them assess if they are correct or misleading. However, this possibly will only test students’ memorisation skills. Increasing students’ understanding will increase their problem-solving and cognitive skills needed to recognize misleading data visualisations (Borthick & Clark, 1986). Essay questions are most suitable to test students’ understanding and if they can apply knowledge in real-life situations (Snyder & Snyder, 2008). Alternatively, to test whether students recognized and understand a breach in graph conventions, the use of language in open questions can be studied. Vocabulary often serves as a proxy for knowledge and can be used to measure the literacy and numeracy of a person (Fisher et al., 2009). Furthermore, in certain disciplines, specific terms (e.g. X-axis and Y-axis) are important for the understanding of a subject, as is the case in the field of mathematics (Shanahan & Shanahan, 2008).

Designing data visualisations

The visual element of graphs can help the general public to more easily interpret data. When designing data visualisations, multiple elements influence the potential of a graph to transfer information to the general public. It is important to pay attention to colour, symbols, dimension, scale, and most of all choosing the right chart type (Brett et al., 2020; Iliinsky & Steele, 2011). However, data visualisations are not purely visual but also involve numerical and textual information that support the visual message. Examples of "visual supports" (Brett et al., 2020) are: legends explaining what variables are present, titles explaining what de visualisation is about, axes labels, values on the axes, data source, and a subtext underneath a figure (Allen, 2021). No strict rules are present for graph design, but design standards were developed over time. Following such graph conventions, including the characteristics described above, increases the effectiveness of a graph and can limit potential misinterpretations.

Misleading data visualisations

Although data visualisations are often designed to objectively (Kennedy et al., 2016) portray data, graphs in news media are always used to frame a specific story. Design choices often help media to support and fortify their argument. When graphs are poorly or maliciously designed, they can become misleading (Lo et al., 2022). Following the definition of Cairo (2019), when I speak about misleading data visualisations, I refer to "graphs that violate common graph conventions to optimize its rhetorical power despite original data, and graphs that lead us to spot patterns and trends that are dubious, spurious, or misleading". According to Cairo (2019), graphs can be misleading due to poor design, by displaying only part of all data, by displaying dubious data, by hiding uncertainty, or by suggesting misleading patterns.

Within the category of poor design, multiple mistakes are commonly made when designing data visualisations. Nguyen et al., (2021) defined four categories in which design pitfalls are most common: colour, shape, size, and spatial orientation. Examples of common violations within these groups include going against colour norm (e.g. using blue for warmer colours instead of red), using the wrong type of graph (e.g. bar chart instead of a histogram), omitting axes (Y-axis does not start at zero or X-axis only shows a selective period or is not evenly spread), and using a 3D-effect which puts more emphasis on items in the front (Kelleher & Wagener, 2011; Nguyen et al., 2021).

Other common data visualisations mistakes as defined by Cairo (2019), such as cherry-picking, can be harder to recognize. It is up to the researcher to show all data, uncertainty in the data, and to be transparent about how data was collected. The vast collection of data visualisations on social media makes it hard if not impossible for the general public to check the integrity of a researcher or certain project. It is expected that students can more easily learn to recognize design-based mistakes in data visualisations compared to checking the background and uncertainty of a visualisation and time for a lesson on misleading data visualisation is limited, as most lessons only last 45 minutes. Therefore, I focused on design-based misleading data visualisations during this project.

Graph education

Cognitive functioning can vary between students of different types of education. Based on educational level, different teaching strategies are needed to educate students on important topics (Lövdén et al., 2020). Although little is known about teaching strategies specifically aiming at increasing graph literacy of students in vocational education programmes (Börner et al., 2019), more general teaching strategies can be used when designing a lesson on data visualisations.

From research on educational videos, we know that students pay less attention to an explanation that is easy or logical as they are not stimulated to think for themselves. As a result, students do not link new knowledge to their prior knowledge (Muller et al., 2008). However, activating prior knowledge is crucial for student learning. By activating background knowledge, educators can guide learning to expand students' knowledge and clarify their misconceptions (Nurpahmi, 2015). To increase students' learning potential, it is therefore important to link to prior knowledge and find the right balance between easy and logical theory and more challenging theory. Additionally, presenting students with alternative conceptions can increase their mental effort to understand a particular subject (Muller et al., 2008).

Another useful and widely applied teaching strategy is inquiry-based learning. Inquiry-based learning is a pedagogical approach for discovering new causal relations, with the learner formulating questions and making observations (Pedaste et al., 2012). Benefits include increased engagement and critical thinking. According to Banchi & Bell (2008), inquiry-based learning can be divided into four different forms: confirmation inquiry (confirming a principle when results are already known), structured inquiry (i.e. investigating teacher-presented questions through a set procedure), guided inquiry (investigating teacher-presented question using student-designed processes) and open inquiry (investigating own questions that are designed through students-designed procedures). Regardless of which form is applied, to use inquiry-based learning to its full potential, it is important to provide adequate guidance to students during learning activities (Lazonder & Harmsen, 2016).

To grab students' attention and keep them engaged with the lesson, my study uses examples of real-life misleading graphs. It is important that the right examples of misleading data visualisations are chosen to keep students engaged with the lesson. Multiple socio-cultural factors, including subject matter, source, beliefs, and opinions, affect engagement with data visualisations (Kennedy, Hill, Allen, et al., 2016). Therefore, when choosing examples of data visualisations for vocational education students, it is essential to investigate their interests prior to designing educational materials.

Besides focusing on students, teachers should be actively included in the process of increasing students' graph literacy by providing them with the right support. This is crucial as including instructional strategies for teachers in curriculum materials can positively influence student learning outcomes (Cervetti et al., 2015; Roblin et al., 2018). Teacher support strategies can be divided into educational supports (i.e. teacher learning) and procedural supports (i.e. assistance with the implementation of a curriculum). Both can be addressed in a one-page teacher manual. To optimize student learning outcomes this manual should, at least, include a description of instructional strategies to support student learning (e.g. class discussion or assignments), extra information on the subject (e.g. additional background information), and information about student ideas (e.g. students prior knowledge or alternative conceptions) (Roblin et al., 2018).

Research questions and hypotheses: increasing graph literacy and preventing misinformation

As vocational education students struggle with interpreting data visualisations and recognizing misleading graphs, it is important to increase their graph literacy. To that end, I designed a short introductory lesson on misleading data visualisations. I then investigated to what extent the developed lesson increased graph literacy of vocational education students based on the theory described above. In addition to the lesson, I developed a teacher manual to support teachers to continue education students on misleading data visualisations. To further optimize this manual, I investigated in what ways the developed manual can be of support when preparing lessons about misleading graphs. The research questions are summarised in Table 1.

Table 1: Research questions

RQ1	To what extent does the developed lesson increase students' graph literacy?
• RQ1.1	To what extent does the developed lesson increase students' ability to recognize misleading data visualisations?
• RQ1.2	To what extent does the developed lesson increase students' ability to describe what characteristics can make a graph misleading?
RQ2	In what ways does the teacher manual help a teacher to continue education on misleading data visualisations?

To investigate the effect of the developed lesson on students' graph literacy (RQ1), I used two surveys. Specifically, I focused on the recognition and understanding of misleading graphs. By means of multiple-choice questions, I investigated to what extent the lesson increased students' ability to recognize misleading data visualisations (RQ1.1). Open questions were used to determine if the graphs were recognized for the right reason. Furthermore, I used the open questions to determine to what extent the lesson increased students' ability to describe what graph characteristics are misleading (RQ1.2).

As discussed above, providing teacher support can increase student learning outcomes. To that end, I designed a two-page teacher manual to enable teachers to continue educating students on misleading data visualisations. With interviews, I investigated how the manual can help teachers to continue education on data visualisations (RQ2).

Definitions:

Data visualisations: images that represent raw qualitative or quantitative data and support communication on the data (Azzam et al., 2013).
Graph literacy: the ability and skill to read and interpret visually represented data to extract information from data visualisations, and the ability to recognize common mistakes in graph design and explain why certain mistakes can cause misinterpretations.
Inquiry-based learning: a process of discovering new causal relations, with the learner formulating questions and making observations (Pedaste et al., 2012).
Misleading data visualisations: graphs that violate common graph conventions to optimize its rhetorical power despite original data, and they lead us to spot patterns and trends that are dubious, spurious, or misleading (Cairo, 2019).
Numeracy: understanding of the real number line, time, ratio concepts, fractions, percentages, probabilities, measurement, and estimation (Reyna et al., 2009).
Teacher support: providing learning opportunities for the teacher and providing them with extra material to support continual education on a subject (Roblin et al., 2018).

Material and Methods

I performed a mixed-method classroom study that involved two in-class surveys and multiple semi-structured interviews. This study serves as a pilot for a bigger project. Data from the surveys was collected using Qualtrics' online survey software. The collected data is stored by Dr.W.Wijnker on a secure server from Leiden University.

Procedure
Data for the project was collected in a two-week period between March 31st and April 14th. In the first week, the lesson on misleading data visualisations was given and survey data was collected. In the second week, teachers of the classes were interviewed about the lesson on data visualisations and draft teacher manual.

To measure the effect of a data visualisation lesson on graph literacy, this study involved two in-class surveys. Both surveys were filled in by the same set of participants at the start and end of a lesson on misleading data visualisations. In the surveys, participants were presented with real-life data visualisations found in mainstream media, that match their interests. The lesson, aimed to increase graph literacy, combined with both surveys, had a duration of approximately 30 minutes. Before the start of the first survey participants were asked for consent to use collected data. No personal data was collected. To ensure answers of the first and second survey could be linked, at the start of both surveys, participants were asked to fill in the participant's number that they received before the start of the lesson. Anonymized data, collected by researchers that held a survey before the lesson on misleading data visualisations, in the same classes used for this study, was used to describe the demographic of participants. Additionally, in that survey, the four-item Short Graph literacy scale of Okan et al. (2019) was used to estimate participants graph literacy. Participants received one point for every correct response, resulting in a graph literacy score between 0 and 4 (0 points for the lowest graph literacy, and 4 points for the highest graph literacy).

After all survey data was collected, teachers of the classes that were included in this study, were interviewed on the lesson on data visualisations and the draft teacher manual that was developed for that specific topic. The interviews were semi-structured, using an interview protocol to guide the conversation. At the start of the interview, participants were asked for consent to use the transcribed interviews for analysis. On average, the interviews took 30 minutes to complete. Interviews were recorded. After transcription, all recordings were deleted.

Participants
For this study, I visited ten classes of vocational education in the Netherlands. As this project was a pilot for a bigger study, the sample size was relatively small, and no power analysis was performed. For pilot studies investigating intervention efficiency, a minimum of 25-30 samples is required (Hertzog, 2008; Johanson & Brooks, 2010). In total, 131 students participated in the lesson on misleading data visualisations. Student present in class during the time of the guest lecture, all had to participate in the lesson, including both surveys. Answers from students that did not give consent were deleted. Furthermore, participants that did not complete one or both of the surveys were excluded from the analysis (i.e. progress score in Qualtrics below 100, if only one of two surveys was completed, or if more than three of the six open questions were unanswered). Last, participants that did not answer the open questions seriously, were also excluded from the analysis (e.g. if participants gave answers unrelated to the survey). This resulted in 75 responses that were considered for further

analysis. As a second part of the study, teachers of the classes I visited, were interviewed. Six teacher interviews were conducted during this study.

Instruments

<u>Student surveys:</u>

I created two surveys to test the effectiveness of the lesson for improving graph literacy, each including three graphs (Appendix 1). Graphs used for each survey are displayed in Figure 1. Both surveys were embedded in the lesson itself to promote active participation of students. At the start of the lesson on data visualisations, students filled in a survey to estimate their base-level graph literacy. In this survey, participants were shown three graphs of which they had to assess if they were misleading or accurate (multiple-choice question) and then were asked to explain what they based their choice on (open question). All three multiple-choice and all three open questions were phrased in the same way: "Is this graph misleading" and "Explain why you think this chart is misleading or not misleading". At the end of the lesson on data visualisations, participants filled in the second survey. This survey was a copy of the first survey with new graphs on the same data visualisation mistakes to measure if students' graph literacy increased.

Graphs used for the surveys (Figure 1) were all found on social media. Graphs used for the surveys were selected based on their readability, subject, and type of mistake. Of the three graphs used for each of the surveys, two were misleading and one was correct. One of the misleading graphs contained a mistake on the Y-axis and the other misleading graph contained a mistake on the X-axis. As pointed out by Nguyen et al. (2021) and Lisnic et al. (2023) mistakes on the X-axis and Y-axis are frequently occurring in scientific publications and social media. New visualisations with the same misleading elements were used for the second survey. The order of the graphs was changed for the second survey to prevent students from recognizing the structure of the first survey. Graphs with English titles and axes were translated to Dutch to make them easier to understand for students, as not all students are fluent in English.

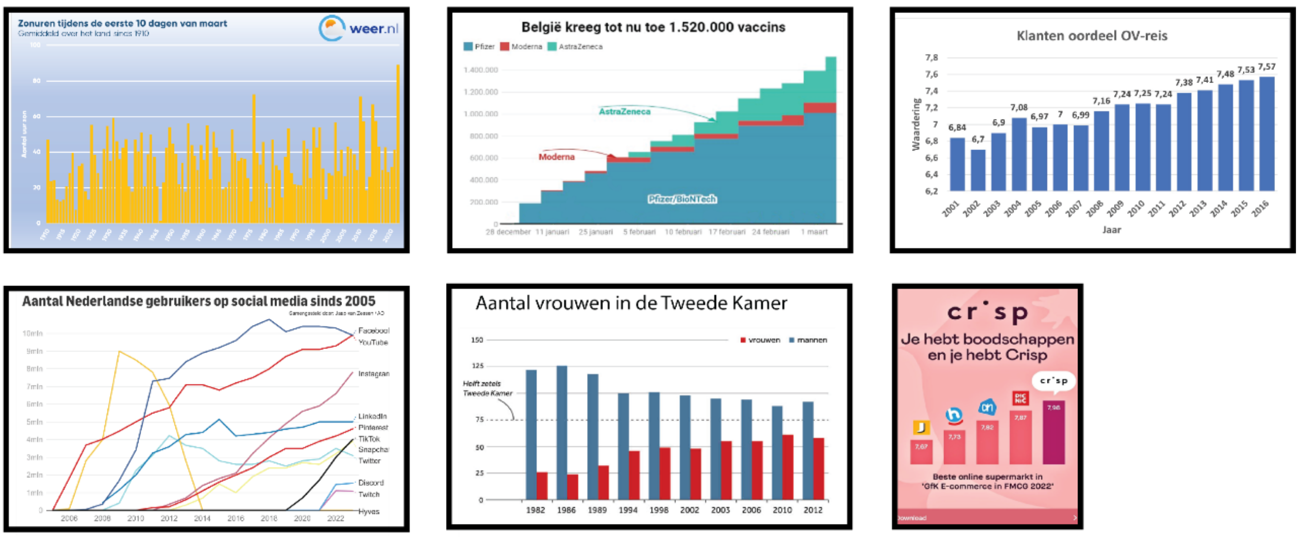


Figure 1: Graphs used for student surveys.

The average time spend answering one survey was two minutes and 40 seconds. All participants answered the same questions in the same order. To optimize results of the study, participants were told there are no 'right' or 'wrong' answers, and that it is important for the study that they do not discuss or copy each other's answers. Furthermore, an incentive was provided in the form of easter eggs to increase the chance that participants answer questions seriously.

Lesson on misleading data visualisations:
Students participated in a lesson on data visualisations developed to increase their graph literacy (e.g. their ability to recognize misleading data visualisations and identify the type of deception). During this 30-minute lesson, the three graphs (i.e. two misleading and one correct graph) from the first survey were discussed in detail. Additionally, more graphs with a mistake on the Y-axis or X-axis were shown to familiarize students with these frequently occurring types of misleading data visualisations. Graphs used in the lesson are depicted in Appendix 2. Following the theory described in the theoretical framework, the lesson used real examples of misleading data visualisations, relevant to students' daily life to keep them engaged with the lesson and trigger critical thinking. Furthermore, the lesson used techniques from confirmation inquiry-based learning as described in the theoretical framework, as this is often used to reinforce prior knowledge.

The lesson started with a short general introduction, carefully chosen to prevent triggering students' graph literacy. Then students answered the first survey. After the survey, the goal and the importance of the lesson were explained. Then, the first misleading graph was discussed in five steps: subject of the graph, type of graph, what is shown on the X-axis, what is shown on the Y-axis, and whether and why the graph is misleading or not. Of each misleading graph a correction was presented to highlight why certain mistakes can be misleading. Other graphs used in the lesson were discussed the same way. The last graph that was discussed was a correct graph, to indicate that not all graphs in the media are misleading. Following the second survey, students were provided with answers to the second survey, and a short summary to conclude the lesson.

Teacher interviews:
During the second week of data collection, teachers were interviewed. The interviews were used to investigate in what ways the teacher manual (Appendix 3) can help a teacher to continue education on misleading data visualisations. Questions on the interview guide (Appendix 4) included questions that discuss the different aspects of the teacher manual and aspects that might be missing from the manual. That way, based on this pilot study, the teacher manual can be improved for the big project on educating about misleading data visualisations.

Data analysis
To test to what extent the lesson on data visualisations increased students' ability to recognize misleading and correct graphs, I quantitatively compared answers from the first survey with answers from the second survey. I analysed answers related to the graph with a mistake on the Y-axis, the graph with a mistake on the X-axis, and the correct graph, separately. Responses that lacked an answer to the open question in the first and/or second survey were removed. The remaining entries were then classified as rightfully recognized (i.e. graph is correctly labelled as correct or misleading with correct accompanying reason) or not recognized (i.e. either incorrectly assigned or without the right substantiation). To determine if there is a significant difference between the number of correct recognitions in the first and second survey, I used the mcNemar test, on above-described classifications. As three tests were performed, the Bonferroni correction was applied.

Following the quantitative analysis, I analysed the open questions of the student surveys to investigate if students use more technical terms to describe the graph in the second survey compared to the first one. By paying attention to the vocabulary used by students to describe presented graphs, I investigated if their vocabulary has changed to using more technical terms which can be an indicator for increased graph literacy.

Last, I transcribed and summarised teacher interviews to determine in what ways the teacher manual can support teachers with educating students about misleading graphs. All interviews were transcribed after which they were summarized. A list was made of all strengths, weaknesses, and possible improvements for a second version of the manual.

Results

Educating students on misleading data visualisations:

Descriptive statistics: demographics

In total, 131 students participated in the lesson on misleading data visualisations. After removing answers without consent, that were incomplete or unserious, 75 remaining responses were used for analysis. Participants were aged between 16-26 (M= 17.9, SD= 2.5). Of the participants, 21 identified as female, 44 identified as male, four identified as other, and three participants did rather not disclose their gender. All participants were in the first year of their vocational education programmes (helping care and well-being, social care, engineering, ICT, software development, media design, content creator, or law enforcement), with levels of education ranging between 2-4. On average, the students had a graph literacy score (scale 0-4) of 2 (SD= 0.9) (Figure 2).

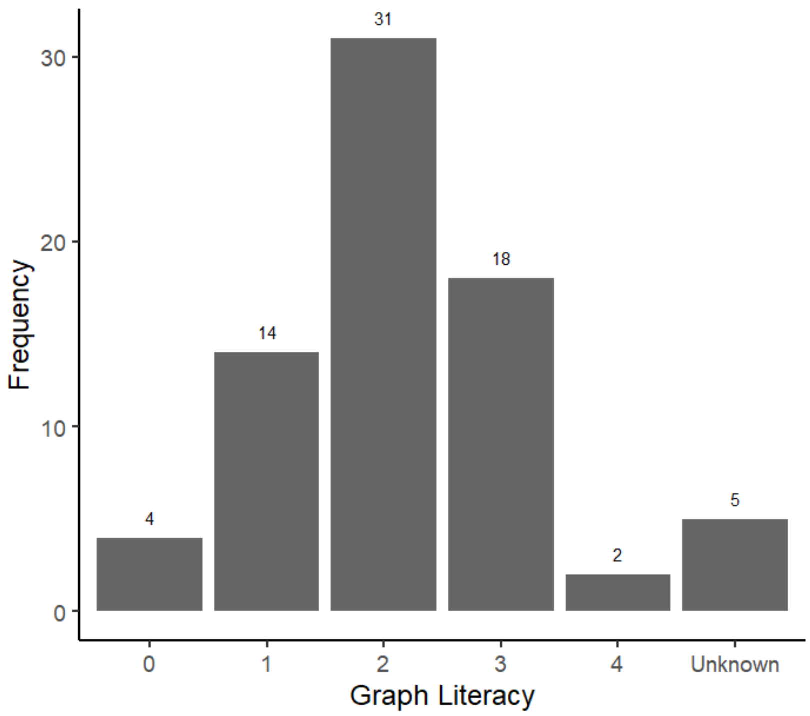


Figure 2: Graph literacy scores of participants. Unknown= not answered or not completed questions of graph literacy test.

Student survey's

RQ 1.1 When investigating students' ability to correctly recognize correct and misleading data visualisations, we looked at their ability to recognize correct graphs, graphs with a mistake on the Y-axis, and graphs with a mistake on the X-axis. As I performed multiple statistical analyses on the same data set, the alpha value was corrected for multiple testing using the Bonferroni correction (α= 0.0167).

Data visualisations with a mistake on the Y-axis:

The first analysis of students' ability to point out misleading data visualisations with a mistake on the Y-axis showed that there is a significant difference ($p < 0.001$) between their ability to do so at the start and end of the lesson on misleading data visualisations (Figure 3). The Cohen's *g* is higher than 0.25, therefore, the effect is large (C_g= 0.36). As visualized in Figure 3, 38 participants did not recognize the mistake in the first survey but did recognize the mistake in the second survey (top left), 22 participants did not recognize the mistake in both the first and second survey (top right), six participants recognized the mistake both in the first and the second survey (bottom left), and six participants recognized the mistake in the first survey but did not recognize the mistake in the second survey (bottom right). This indicates that the lesson can improve students' ability to recognize misleading data visualisations containing a mistake on the Y-axis.

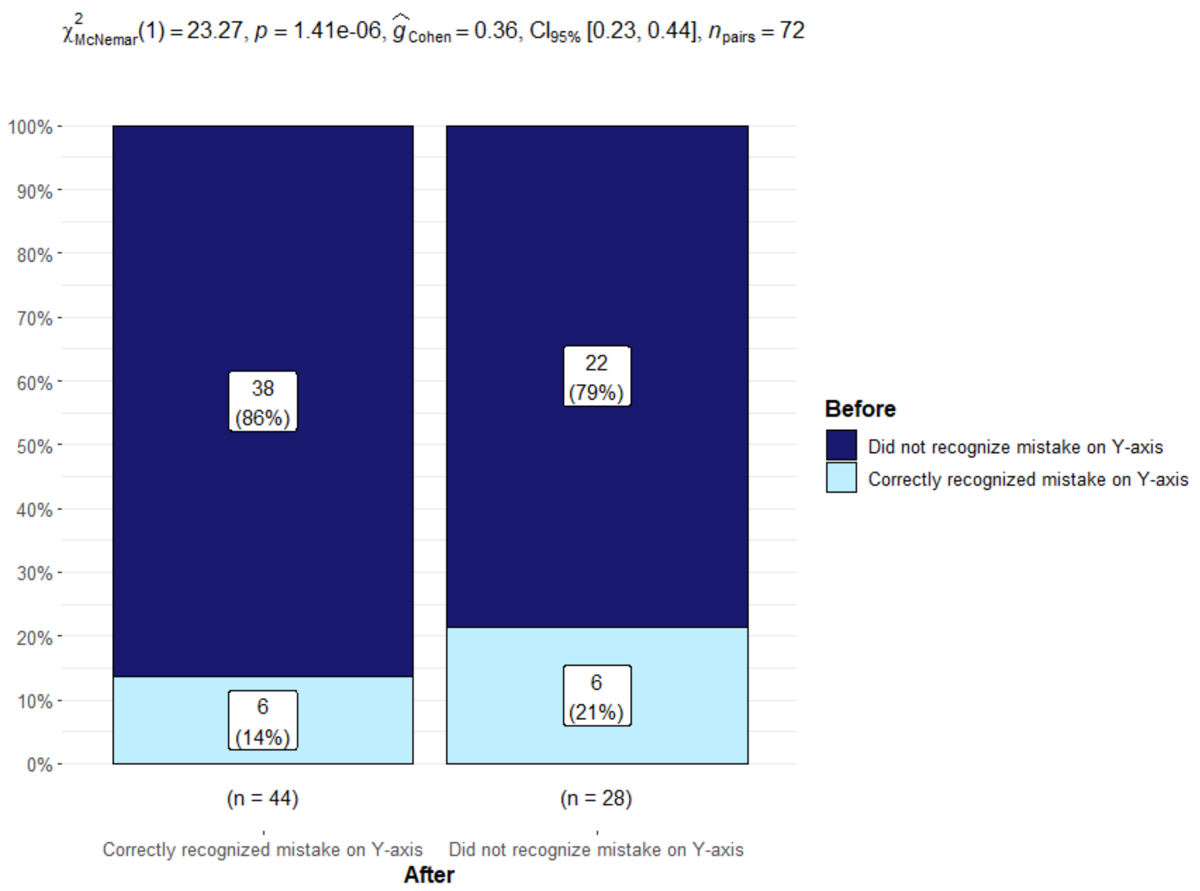


Figure 3: Classification of students' ability to correctly recognize a misleading graph with a mistake on the Y-axis before and after the lesson on misleading data visualisations.

However, after the data collection, it became apparent that the graph used for the second survey (Figure 1, bottom right) contained both a mistake on the Y-axis and on the X-axis. Upon further analysis of the answers to the open questions, it seemed that some students only recognized the mistake on the X-axis instead of the mistake on the Y-axis, or both. Therefore, the previous test only shows that there is a significant difference in their ability to recognize a misleading graph in general and does not disclose anything about their ability to specifically recognize a mistake on the Y-axis. When answers without clear recognition of the mistake on the Y-axis were classified as not recognized instead of rightfully recognized, there was no significant difference ($p = 0.0186$) between students' ability to recognize a mistake on the Y-axis at the start and end of the lesson (Figure 4). The effect was medium ($C_g = 0.23$). To make sure this adapted classification was reliable, a second coder analysed the answers to the open questions. Coding of the second coder was in accordance with the coding of the first coder. As is visible in Figure 4, 19 participants did not recognize the mistake on the Y-axis in the first survey but did recognize the mistake in the second survey (top left), 41 participants did not recognize the mistake on the Y-axis in both the first and second survey (top right), five participants recognized the mistake both in the first and the second survey (bottom left), and seven participants recognized the mistake in the first survey but did not recognize the mistake in the second survey (bottom right). This suggests that although the difference is not significant, nineteen students did improve their ability to recognize misleading graphs with a mistake on the Y-axis.

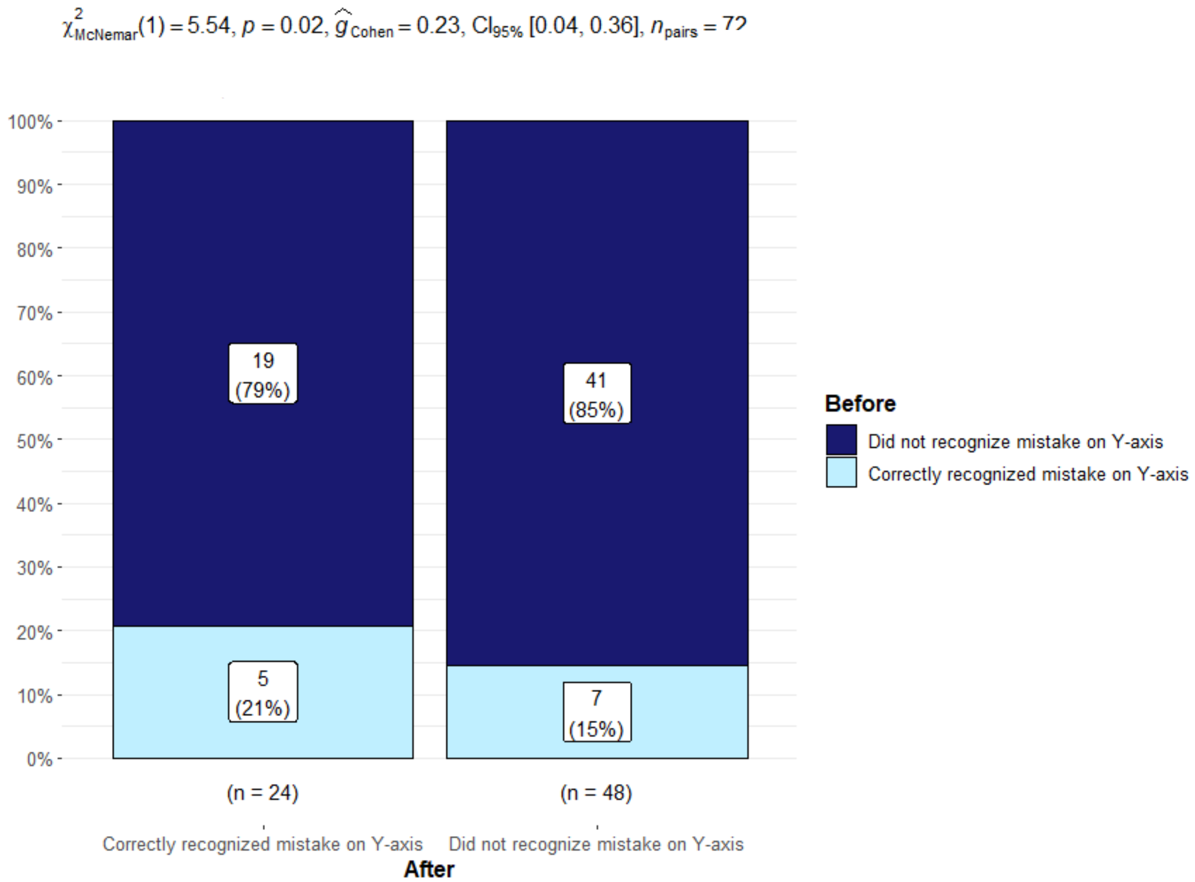


Figure 4: Classification of students' ability to correctly recognize a misleading graph with a mistake on the Y-axis before and after the lesson on misleading data visualisations, corrected with answers to the accompanying open question.

Data visualisations with a mistake on the X-axis:

Analysis of students' ability to point out misleading data visualisations with a mistake on the X-axis shows that there is a significant difference between their ability to do so at the start and end of the lesson on misleading data visualisations (Figure 5). The Cohen's g is higher than 0.25, therefore, the effect is large (C_g= 0.50). As is visible in Figure 5, 36 participants did not recognize the mistake in the first survey but did recognize the mistake in the second survey (top left), 33 participants did not recognize the mistake in both the first and second survey (top right), and one participant people recognized the mistake both in the first and the second survey (bottom left). This demonstrates that over half of the students increased their ability to recognize misleading graphs with a mistake on the X-axis.

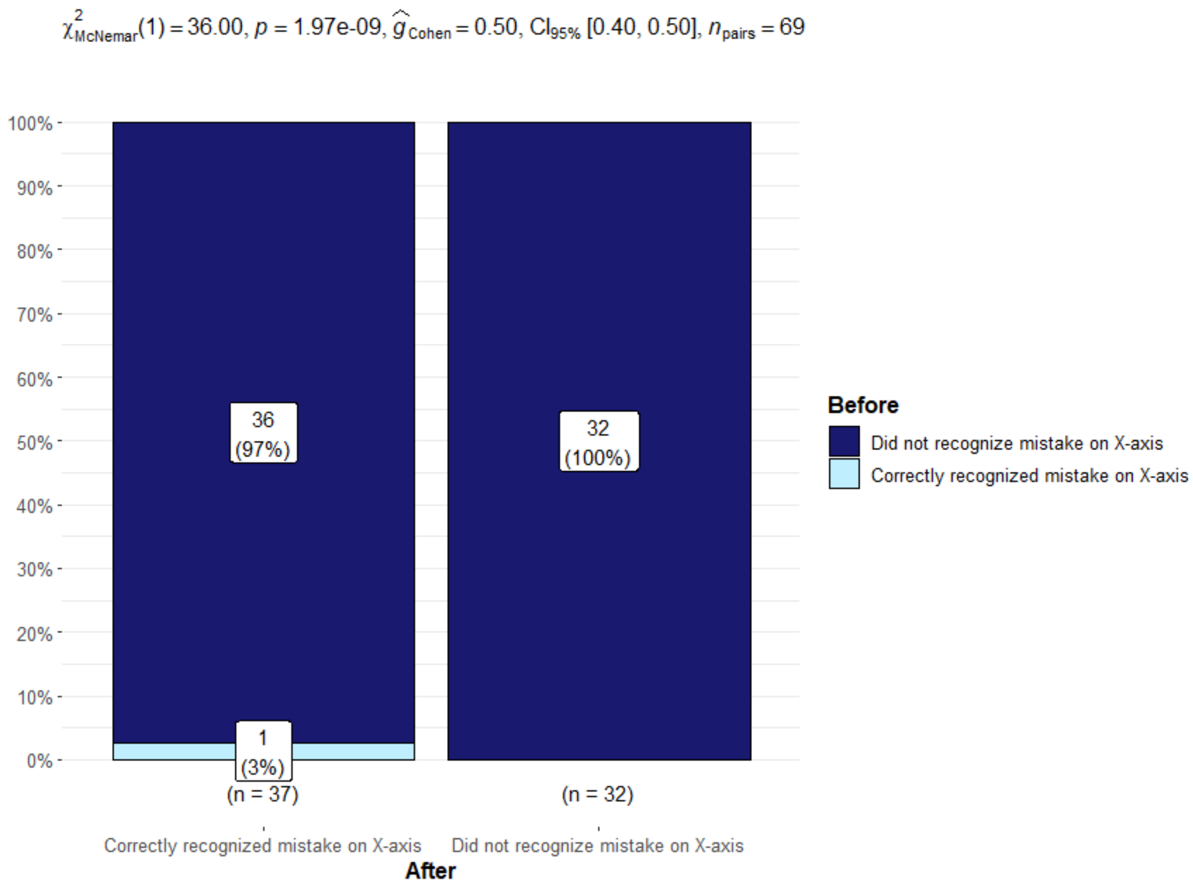


Figure 5: Classification of students' ability to correctly recognize a misleading graph with a mistake on the X-axis before and after the lesson on misleading data visualisations.

Correct data visualisations:

Analysis of students' ability to point out a correct graph between two misleading ones, showed that there is no significant difference (p= 0.56) between their ability to do so at the start and end of the lesson on misleading data visualisations (Figure 6). As is visible in Figure 6, fifteen participants did not recognize the correct graph in the first survey but did recognize it in the second survey (top left), 21 participants, in both surveys, did not recognize that the graph was correct (top right), 23 participants recognized the correct graph in both surveys (bottom left), and twelve participants recognized the graph in the first survey but did not recognize it in the second survey (bottom right). Although some students did improve and recognized the mistake in the second survey, this suggests that students struggle to recognize a correct graph in between two misleading graphs.

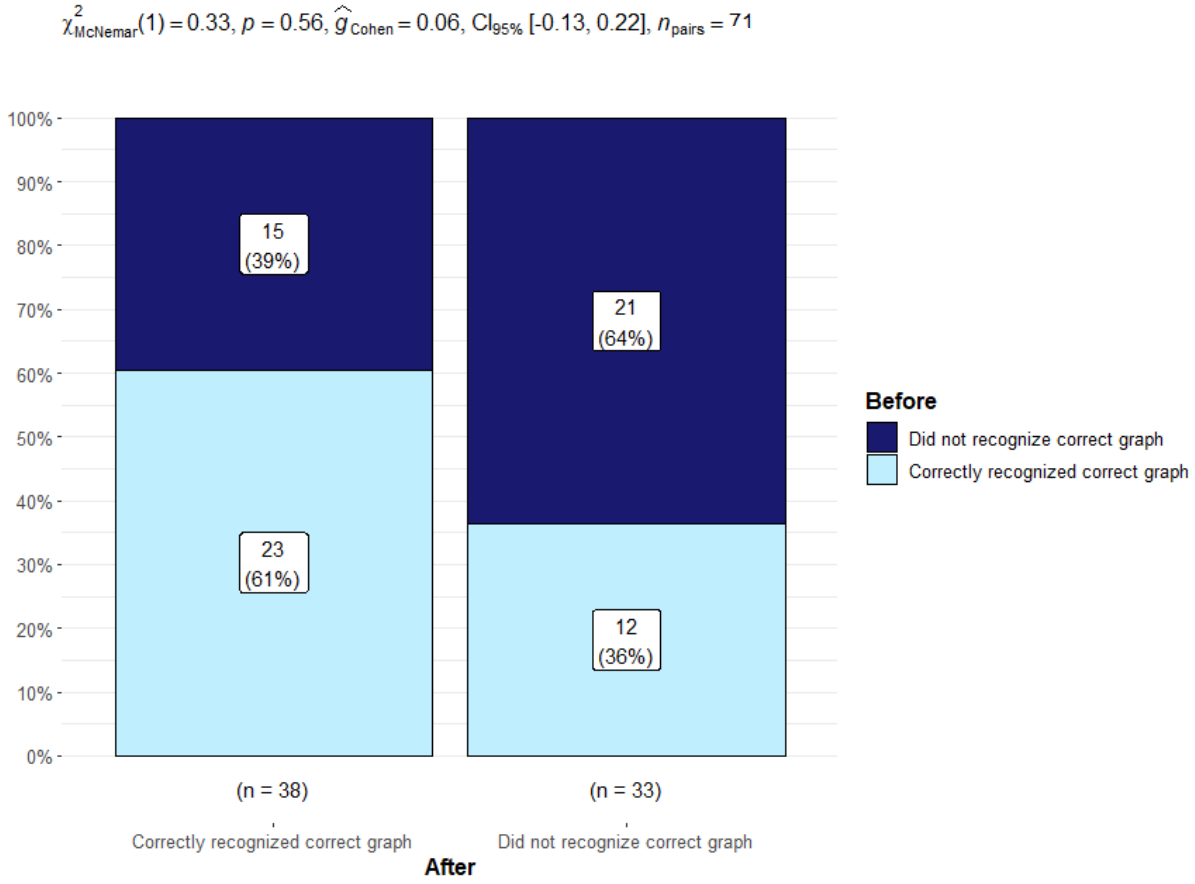


Figure 6: Classification of students' ability to correctly recognize a correct graph before and after the lesson on misleading data visualisations.

RQ 1.2 To investigate to what extent the developed lesson increases students' ability to describe what characteristics can make a graph misleading, I analysed the use of vocabulary in the open-ended questions of the surveys. In the first survey, the terms X-axis and/or Y-axis were not named a single time. In the second survey the terms X-axis and/or Y-axis were named 34 times, by 24 different participants. This suggests that the lesson about misleading data visualisations increased students' ability to describe what characteristics can make a graph misleading.

Teacher manual misleading data visualisations:

Descriptive statistics: demographics

In total, six teachers from different locations of MBO Amsterdam/Flevoland were interviewed for this study. Age of participants varied between 31 and 65, two identified as male, and four identified as female. Background education varied between participants (i.e. pedagogy, primary school teacher, Dutch, painting techniques, law, or mechanical engineering). All teachers had 1.5 to 15 years' experience with educating mathematics at the time of the interview. Every teacher had experience with teacher manuals, but they do not use the manuals with equal frequency.

Teacher interviews:

RQ 2 Semi-structured interviews were used to investigate in what ways the developed teacher manual can help teachers with education on misleading data visualisations. The interviews were summarized. Suggestions for improvements are provided in Table 2.

In general, all teachers indicated that graphs are part of the current exam program, but that misleading graphs are not discussed. Although teachers indicated that misleading graphs are not part of the program, they think it is very important to discuss misleading data visualisations with students in class. As misleading graphs are not part of the standard program, all six teachers indicated that a teacher manual can be of great help when preparing mathematics lessons on misleading data visualisations. However, their evaluation of the format of a manual differed. Three teachers indicated that the developed manual provided little support as it was too short. They preferred a more detailed, textual manual with examples of presentations and assignments, more like a syllabus. For them, the manual did not provide enough background information. Two of them suggested adding a short explanation to the tools and core concept boxes as they did not have the knowledge to explain these concepts by themselves. The three other teachers indicated that the manual was of great help when preparing a lesson on misleading data visualisations as it provides a short overview of important theory and possible work forms. They indicated that having a short overview of the most important theory and concepts is especially helpful when having little time to prepare lessons. A detailed syllabus would take too long to read. In general, more experienced math teachers (1.5-5 years of experience) tended to prefer the short manual, while less experienced teachers (10-15 years of experience) preferred a more detailed guide. Boxes with the learning goals and possible work forms were experienced as useful when preparing lessons by all six teachers. Two teachers asked to add methods for checking if the learning goals are achieved, as it is sometimes difficult to estimate if a lesson is too difficult, because the background knowledge of students varies. Three teachers indicated that the manual not only helps to prepare a lesson, but it also adds to their own knowledge.

Table 2: Suggestions for improving the teacher manual as indicated by teachers during the interviews.

Suggestions for improvement
Move learning goals to the top of the manual
Add steps for interpreting graphs used during the lesson
Add a time indication for work forms
Add possible methods to check if learning goals are achieved
Add a short explanation of tools and core concepts
More clearly indicate that the manual can be used to develop multiple lessons
Break down learning goals into more specific goals

Overall, all teachers were satisfied with the manual. Even though the manual was not extensive enough for everyone, they all indicated that the manual was useful to have when preparing a lesson. The manual provides guidance and helps to save time. Suggestions were used to update the manual.

Discussion

General discussion:

Data visualisations are powerful tools to communicate to the general public about complex data. The visual element helps people to correctly understand complicated matter (B. Lee et al., 2020). However, sometimes the visual elements of graphs are misused; they can accidentally or purposefully be used to frame a story and change people's attitude towards a subject (Lo et al., 2022; Pandey et al., 2014). That way misinformation can be spread. I set up a project to investigate how to increase graph literacy of vocational education students to prevent misinformation by misleading graphs from spreading. I developed a lesson on misleading data visualisations, where I focussed on educating students on frequently occurring mistakes in data visualisations. Results indicated that a short lesson on misleading data visualisations can improve students' ability to recognize misleading graphs with a mistake on the X-axis and to a lesser extent increase their ability to recognize misleading graphs with a mistake on the Y-axis. Of the 77 participants of this study, 36 participants improved their ability to recognize an error on the X-axis (Figure 5), and 19 participants improved their ability to recognize a mistake on the Y-axis (Figure 4). Recognizing a correct graph amongst misleading graphs proved to be difficult for students, even after the lesson on misleading data visualisations.

For the student surveys, data visualisations found in mainstream media were used. After data collection, it became apparent that the graph used to assess students' ability to recognize a mistake on the Y-axis in the second survey (Figure 1, bottom right), also contained a mistake on the X-axis. Students pointed out either the intended mistake on the Y-axis or the mistake on the X-axis, but not both. Presumably, students stopped looking for mistakes once they recognized one of the two mistakes. This may have caused an underestimation of students' ability to recognize a mistake on the Y-axis after the lesson, which may explain why I did not measure a significant improvement in their ability to do so (Figure 4). Yet, even without accounting for this possible underestimation, 19 participants did improve their ability to recognize misleading graphs with a mistake on the Y-axis. Even though this may not be significant, this is a relevant improvement as every person able to correctly interpret a data visualisation helps to stop misinformation from spreading. When not taking into account which mistake was recognized and labelling both mistakes as correct, there was a significant difference in students' ability to recognize misleading graphs before and after the lesson (Figure 3). This shows that, although not all errors were pointed out, students' overall ability to recognize misleading data visualisations did increase. To avoid confusion among students in future lessons, the graph with a mistake both on the X-axis and Y-axis, should be replaced by a graph with a single mistake on the Y-axis.

Contrary to results described above, there was no effect of the lesson on students' ability to recognize correct data visualisations (Figure 6). Answers to the open questions from the survey suggest that, even after the lesson including examples of correct graphs, students still struggle to recognize correct data visualisations. Roughly a quarter of the participants pointed out that they struggled to understand or interpret the correct graph used in the second survey. One of the students said: "I am unable to see what is happening in the graph, it's too chaotic". Another student said: "Too much is happening". These results suggest that reading graphs in general can be difficult for some students and more education is needed to improve their ability to read and interpret graphs. To ensure that all students can understand the introductory lesson, difficult graphs should be replaced by more unambiguous graphs in the final design of the lesson.

In addition to the lesson on misleading data visualisations, I developed a teacher manual to support teachers to continue educating students on misleading data visualisations. I forwarded the manual and during separate interviews discussed the strengths and weaknesses of the manual with the six vocational education teachers of the classes visited for the guest lecture. They indicated that the developed manual is useful when preparing lessons as it saves time and provides background knowledge that they did not have ready themselves. Learning goals, core concepts, background knowledge of students, and links to extra information were experienced as helpful to have when preparing a lesson in a limited amount of time. This is in line with the findings of Roblin et al. (2018). However, teachers were divided on the length of the manual, half wanted a more extensive manual including complete lessons, and the other half was satisfied with the short manual as they indicated that a short manual is easier to quickly read through before a lesson. Generally, teachers with more experience in teaching math preferred the short manual, while teachers with less experience in mathematics would prefer a more detailed manual. Even though some teachers would rather have had a more extensive manual including materials, all six indicated that the developed manual could be of help when preparing a lesson in misleading data visualisations. Thus, a teacher manual is experienced as very helpful when preparing lessons. Depending on the background knowledge of the teacher, additional support should be provided to help prepare lessons on misleading data visualisations. When using the lesson developed for this study, directions for using the presentation must be followed to achieve the results as described in this report.

Limitations and future research:

In this study, I investigated the effectiveness of a short lesson about misleading graphs on graph literacy of vocational education students. Results were positive. However, before students participated in the lesson and answered both surveys, they participated in another survey for a different project. In that survey, students were asked to evaluate the size difference between two groups in misleading graphs and their corrections. Reading and evaluating graphs could have triggered students' graph literacy. This could have resulted in an underestimation of the results of this study, as some students potentially already recognized the misleading graphs in the first survey because their graph literacy was triggered. Additionally, the survey prior to the lesson could have influenced students' attention span and willingness to participate in another survey. For future projects, it is recommended not to conduct two studies in the same class on the same day. That way, no prior knowledge is triggered in advance.

By triggering students' interest in a subject, educational opportunities can be improved (Renninger & Hidi, 2020). To optimize the learning potential of the lesson, we tried to match graphs used in the lesson and student surveys to students' interests. However, all graphs used were sourced from social media, limiting the choice of graphs to use, meaning that the lesson was not optimally matched to students' interests. To further increase the learning potential of the lesson, alternating self-made graphs with graphs sourced on social media, could be a solution to better match graphs to the interests of the target audience and enhance the learning potential. Caution should be paid when using self-made graphs, as the urgency and importance of the lesson can become less clear to students when little or no real-life examples are used.

Results indicated that students struggled to read and interpret some of the graphs used in the surveys. Although basic terms (e.g. X-axis, Y-axis, and bar chart) were explained during the lesson, students did not gain enough skills to interpret more complicated graphs as for example the correct graph used for the second survey (Figure 1, bottom right). For future studies, it is therefore recommended to first give an introductory lesson about reading and interpreting data visualisations, before educating students on misleading data visualisations. That way, when discussing misleading data visualisations, all students will have the same background knowledge required to interpret graphs, allowing them to focus on recognizing mistakes in data visualisations.

No follow-up survey was included in the design of this study. Therefore, it is currently unclear what the long-term effect of the lesson is, on students' graph literacy. To prevent misinformation by misleading graphs from spreading, it is important that the effect of the lesson is long-lasting. A more extensive study, including one or multiple follow-up surveys, could shed light on the sustainability of the effect over time and might point out whether additional lessons are needed to increase students' graph literacy on the long term.

This study aimed to determine whether graph literacy of vocational education students can be improved through education. Results indicate that this is an effective strategy to prevent misinformation from spreading. However, previous studies have shown that media literacy lessons have a weaker effect than fact-checks and corrections (Walter & Murphy, 2018). Combining education and applying fact-checks to misinformation is most effective (Hameleers, 2022). Caution should be paid when educating students on misinformation, because mentioning that misinformation occurs frequently can lower confidence in reliable news sources (Modirrousta-Galian & Higham, 2023).

Conclusion and recommendations

The findings of this study suggest that a short introductory lesson can to some extent increase graph literacy of vocational education students. Reading and correctly interpreting data visualisations remains difficult and more education regarding data visualisations is needed to prevent misinformation by misleading graphs from spreading. A teacher manual can support teachers with preparing these kinds of lessons.

Misleading graphs are everywhere. As long as there are no strict guidelines on the design of data visualisations and misleading data visualisations keep appearing in mainstream media, it is important to educate people about misleading graphs to prevent misinformation from spreading. I recommend all vocational education programmes to introduce thematic lessons on misleading data visualisations and to repeat these lessons over the years students are studying at these programs. I also encourage readers of this report to make people in their surroundings aware of errors they encounter in data visualisations in the media. That way everyone will get the opportunity to learn to take a critical attitude towards graphs in the media. Together we can stop misinformation from spreading.

Acknowledgements

This research was supported by Leiden University Fund (http://www.luf.nl , LUF Lustrum Grant 2020, W20719-1-LLS). I thank Dr.W.Wijnker for the guidance and advice during this project, Sanne Willems and Lonneke Boels for additional advice on teaching and misleading data visualisations, and teachers of MBO Amsterdam/Flevoland for the opportunity to collect data during their class hours.

References

Allen, W. L. (2021). The conventions and politics of migration data visualizations. *New Media & Society*, 146144482110193. https://doi.org/10.1177/14614448211019300

Azzam, T., Evergreen, S., Germuth, A. A., & Kistler, S. J. (2013). Data visualization and evaluation. *New Directions for Evaluation*, *2013*(139), 7–32. https://doi.org/10.1002/ev.20065

Banchi, H., & Bell, R. (2008). The many levels of inquiry. www.teachersource.com

Börner, K., Bueckle, A., & Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, *116*(6), 1857–1864. https://doi.org/10.1073/pnas.1807180116

Borthick, A. F., & Clark, R. L. (1986). The role of productive thinking in affecting student learning with microcomputers in accounting education. *The Accounting Review*, *61*(1), 143–157.

Brett, G., Silver, D., & Beelen, K. (2020). The right tool for the job: problems and solutions in visualizing sociological theory. *Journal for the Theory of Social Behaviour*, *50*(2), 223–248. https://doi.org/10.1111/jtsb.12242

Cairo, A. (2019). How charts lie: getting smarter about visual information (First edition.).

Cervetti, G. N., Kulikowich, J. M., & Bravo, M. A. (2015). The effects of educative curriculum materials on teachers' Use of instructional strategies for English language learners in science and on student learning. *Contemporary Educational Psychology*, *40*, 86–98. https://doi.org/10.1016/j.cedpsych.2014.10.005

Fisher, D., Frey, N., & Leu, D. J. (2009). Background knowledge. Heinemann.

Galesic, M., & Garcia-Retamero, R. (2011). Graph literacy. *Medical Decision Making*, *31*(3), 444–457. https://doi.org/10.1177/0272989X10373805

Hameleers, M. (2022). Separating truth from lies: comparing the effects of news media literacy interventions and fact-checkers in response to political misinformation in the US and Netherlands. *Information, Communication & Society*, *25*(1), 110–126. https://doi.org/10.1080/1369118X.2020.1764603

Hertzog, M. A. (2008). Considerations in determining sample size for pilot studies. *Research in Nursing & Health*, *31*(2), 180–191. https://doi.org/10.1002/nur.20247

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

Iliinsky, N., & Steele, J. (2011). Designing data visualizations. In *Practitioner*.

Jayasinghe, R., Ranasinghe, S., Jayarajah, U., & Seneviratne, S. (2020). Quality of online information for the general public on COVID-19. *Patient Education and Counseling*, *103*(12), 2594–2597. https://doi.org/10.1016/j.pec.2020.08.001

Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. *Educational and Psychological Measurement*, *70*(3), 394–400. https://doi.org/10.1177/0013164409355692

Kelleher, C., & Wagener, T. (2011). Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling and Software*, *26*(6), 822–827. https://doi.org/10.1016/j.envsoft.2010.12.006

Kennedy, H., Hill, R. L., Aiello, G., & Allen, W. (2016). The work that visualisation conventions do. *Information Communication and Society*, *19*(6), 715–735. https://doi.org/10.1080/1369118X.2016.1153126

Kennedy, H., Hill, R. L., Allen, W., & Kirk, A. (2016). Engaging with (big) data visualizations: Factors that affect engagement and resulting new definitions of effectiveness. *First Monday*. https://doi.org/10.5210/fm.v21i11.6389

Lazonder, A. W., & Harmsen, R. (2016). Meta-analysis of inquiry-based learning. *Review of Educational Research*, *86*(3), 681–718. https://doi.org/10.3102/0034654315627366

Lee, B., Choe, E. K., Isenberg, P., Marriott, K., & Stasko, J. (2020). Reaching broader audiences with data visualization. *IEEE Computer Graphics and Applications*, *40*(2), 82–90. https://doi.org/10.1109/MCG.2020.2968244

Lee, S., Kim, S.-H., & Kwon, B. C. (2017). VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics*, *23*(1), 551–560. https://doi.org/10.1109/TVCG.2016.2598920

Lisnic, M., Polychronis, C., Lex, A., & Kogan, M. (2023). Misleading beyond visual tricks: How people actually lie with charts. https://doi.org/10.1145/3544548.3580910

Lo, L. Y., Gupta, A., Shigyo, K., Wu, A., Bertini, E., & Qu, H. (2022). Misinformed by visualization: What do we learn from misinformative visualizations? *Computer Graphics Forum*, *41*(3), 515–525. https://doi.org/10.1111/cgf.14559

Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and cognitive functioning across the life span. *Psychological Science in the Public Interest : A Journal of the American Psychological Society*, *21*(1), 6–41. https://doi.org/10.1177/1529100620920576

Modirrousta-Galian, A., & Higham, P. A. (2023). Gamified inoculation interventions do not improve discrimination between true and fake news: Reanalyzing existing research with receiver operating characteristic analysis. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0001395

Muller, D. A., Sharma, M. D., & Reimann, P. (2008). Raising cognitive load with linear multimedia to promote conceptual change. *Science Education*, *92*(2), 278–296. https://doi.org/10.1002/sce.20244

Nayak, J. G., Hartzler, A. L., Macleod, L. C., Izard, J. P., Dalkin, B. M., & Gore, J. L. (2016). Relevance of graph literacy in the development of patient-centered communication tools. *Patient Education and Counseling*, *99*(3), 448–454. https://doi.org/10.1016/j.pec.2015.09.009

Nguyen, V. T., Jung, K., & Gupta, V. (2021). Examining data visualization pitfalls in scientific publications. *Visual Computing for Industry, Biomedicine, and Art*, *4*(1), 27. https://doi.org/10.1186/s42492-021-00092-y

Nurpahmi, S. (2015). Improving listening skill by activating students' prior knowledge.. Eternal *(English, Teaching, Learning and Research Journal)*, *1*(1), 28–38. https://doi.org/10.24252/Eternal.V11.2015.A3

Oecd,. (2013). Netherlands-country note-survey of adult skills, first results.

Okan, Y., Janssen, E., Galesic, M., & Waters, E. A. (2019). Using the Short Graph Literacy Scale to Predict Precursors of Health Behavior Change. *Medical Decision Making*, *39*(3), 183–195. https://doi.org/10.1177/0272989X19829728

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

Pedaste, M., Mäeots, M., Leijen, Ä., & Sarapuu, T. (2012). Improving students' inquiry skills through reflection and self-regulation scaffolds. *Technology, Instruction, Cognition and learning,* https://www.researchgate.net/publication/285309266

Renninger, K. A., & Hidi, S. E. (2020). To Level the Playing Field, Develop Interest. *Policy Insights from the Behavioral and Brain Sciences*, *7*(1), 10–18. https://doi.org/10.1177/2372732219864705

Research centrum voor onderwijs en arbeidsmarkt. (2019). *Laaggecijferdheid-in-Nederland-Een-overzicht-van-nationale-en-internationale-literatuur*.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin*, *135*(6), 943–973. https://doi.org/10.1037/a0017327

Roblin, N. P., Schunn, C., & McKenney, S. (2018). What are critical features of science curriculum materials that impact student and teacher outcomes? *Science Education*, *102*(2), 260–282. https://doi.org/10.1002/sce.21328

Romano, A., Sotis, C., Dominioni, G., & Guidi, S. (2020). The scale of COVID-19 graphs affects understanding, attitudes, and policy preferences. *Health Economics (United Kingdom)*, *29*(11), 1482–1494. https://doi.org/10.1002/hec.4143

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: Implications for instruction. In *Educational Psychology Review* (Vol. 14, Issue 1).

Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, *78*(1), 40–59. https://doi.org/10.17763/haer.78.1.v62444321p602101

Snyder, L. G., & Snyder, M. J. (2008). Optional teaching, critical thinking and problem-solving skills. *The Delta Pi Epsilon Journal*, 90–99.

van der Linden, S. (2022). Misinformation: susceptibility, spread, and interventions to immunize the public. *Nature Medicine*, *28*(3), 460–467. https://doi.org/10.1038/s41591-022-01713-6

Walter, N., & Murphy, S. T. (2018). How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, *85*(3), 423–441. https://doi.org/10.1080/03637751.2018.1467564

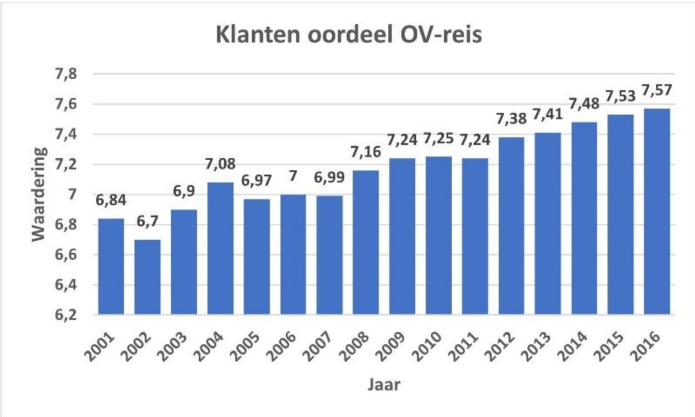
Yang, B. W., Vargas Restrepo, C., Stanley, M. L., & Marsh, E. J. (2021). Truncating bar graphs persistently misleads viewers. *Journal of Applied Research in Memory and Cognition*, *10*(2), 298–311. https://doi.org/10.1016/j.jarmac.2020.10.002

Appendix

1. Student surveys (In Dutch)

1.1 Student survey 1 (at start of lesson)

DC Vul hieronder je deelnemerscode in:

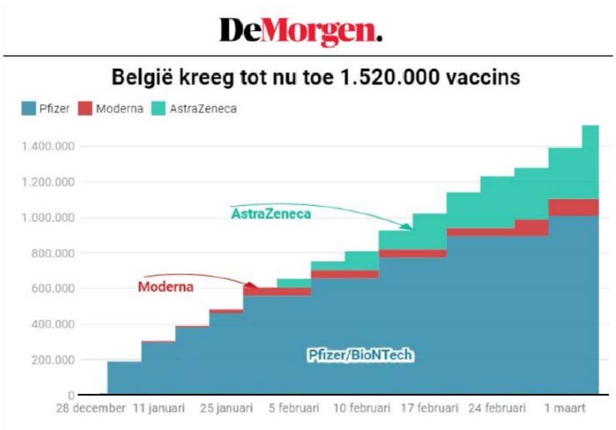


V1.1 Is deze grafiek misleidend?

◯ Ja (1)

◯ Nee (2)

V1.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:

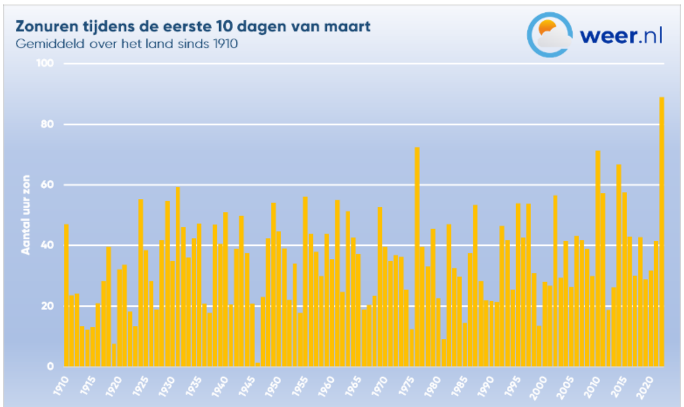


V2.1 Is deze grafiek misleidend?

◯ Ja (1)

◯ Nee (2)

V2.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:



V3.1 Is deze grafiek misleidend?

◯ Ja (1)

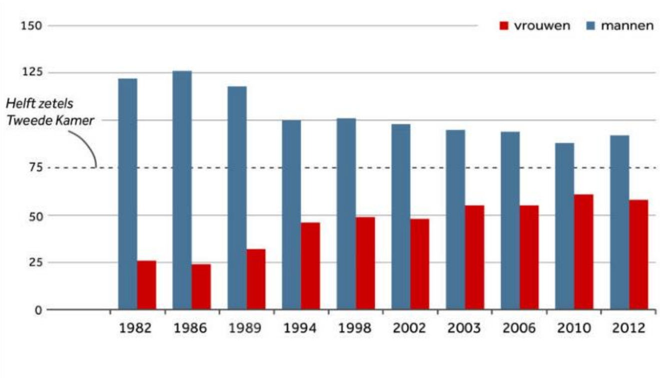
◯ Nee (2)

V3.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:

1.2 Student survey 2 (at end of lesson)

DC Vul hieronder je deelnemerscode in:

Aantal vrouwen in de tweede kamer

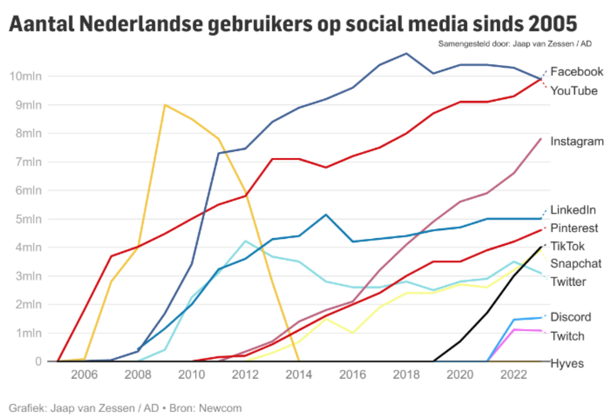


V1.1 Is deze grafiek misleidend?

◯ Ja (1)

◯ Nee (2)

V1.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:



V2.1 Is deze grafiek misleidend?

◯ Ja (1)

◯ Nee (2)

V2.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:

____

V3.1 Is deze grafiek misleidend?

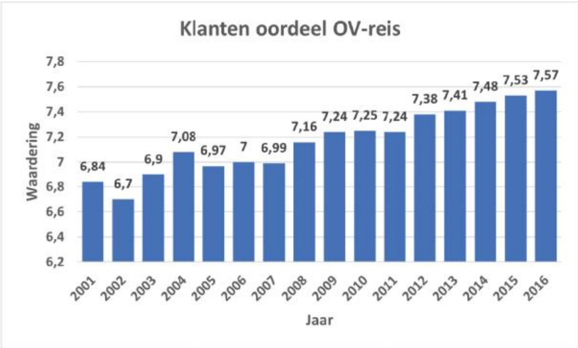
◯ Ja (1)

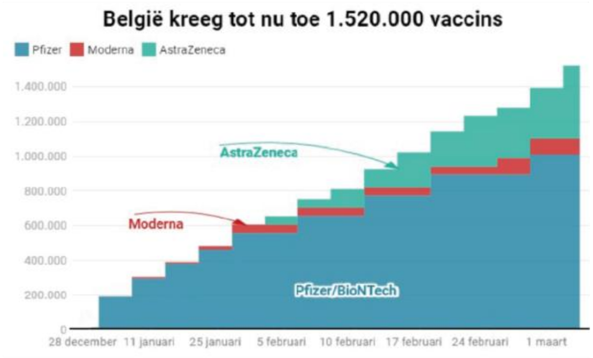
◯ Nee (2)

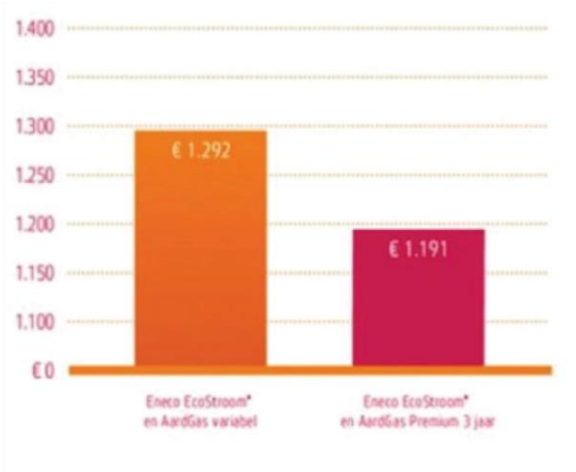
V3.2 Leg uit waarom je denkt dat deze grafiek misleidend of juist niet misleidend is:

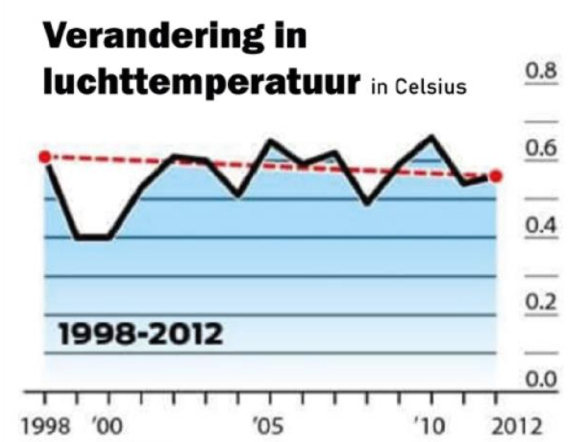
2. Graphs used in lesson on misleading data visualisations

Misleading graphs



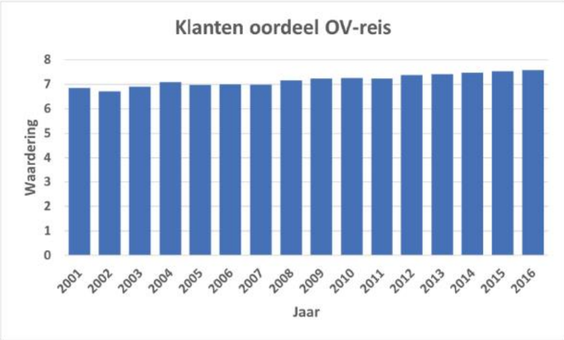


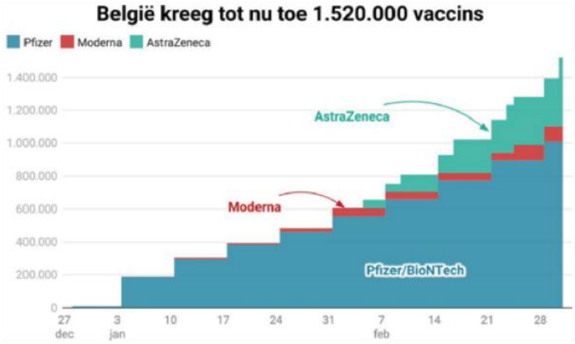


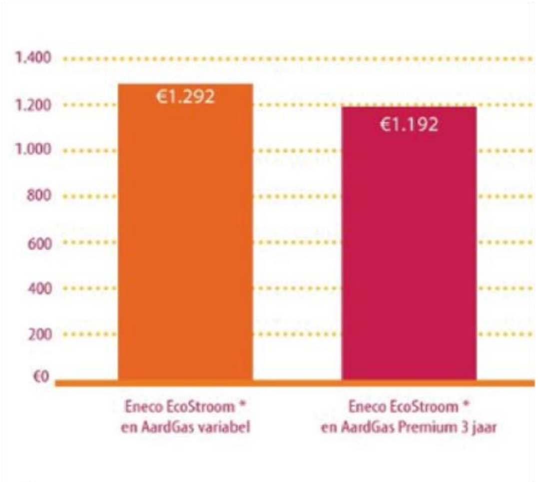


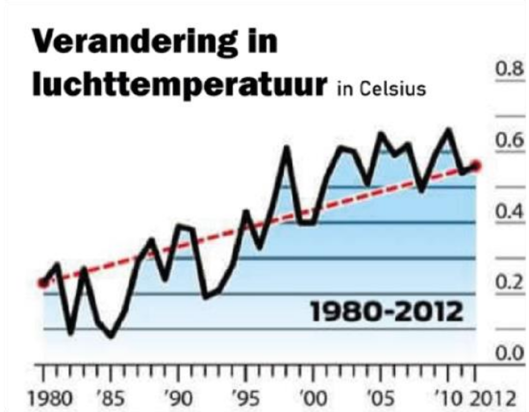
s

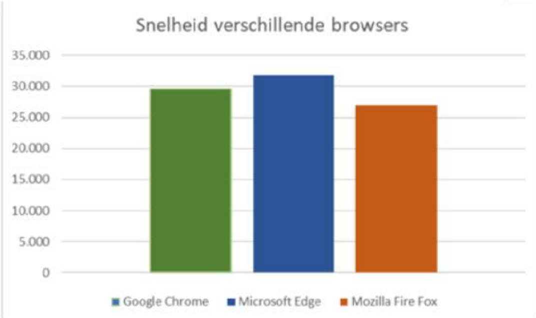
Corrections of misleading graphs

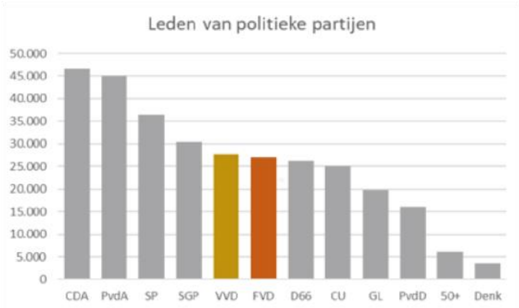




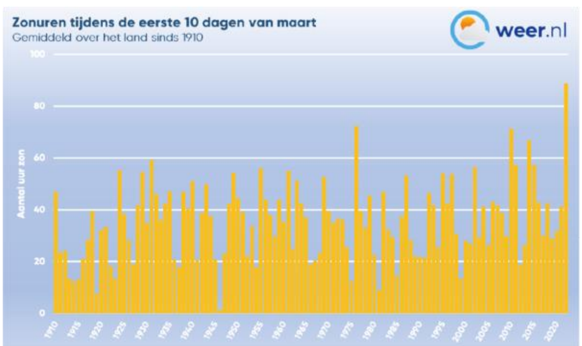








Correct graph



3. Teacher manual (in Dutch)

Docentenhandleiding Misleidende Datavisualisaties

Datavisualisaties worden vaak gebruikt om complexe informatie over te dragen aan het algemene publiek. Hoewel grafieken gegevens objectief kunnen weergeven, worden deze in sommige media gebruikt om een verhaal naar eigen hand te zetten. Dit kan, doelbewust of per ongeluk, leiden tot verkeerde interpretaties van getoonde gegevens. Om jongvolwassenen voor te bereiden op onze snel digitaliserende wereld, waarin ze dagelijks grafieken tegen komen, is het van belang om ze te onderwijzen in het herkennen van misleidende datavisualisaties. Deze docentenhandleiding bevat extra informatie over (misleidende) datavisualisaties en tips en handvatten voor het geven van lessen over misleidende datavisualisaties.

Hulpmiddelen bij het visualiseren van data:

- Histogrammen
- Circeldiagrammen
- Staafdiagrammen
- Lijndiagrammen
- X-as
- Y-as
- Schaal

Achtergrondkennis:

Na het afronden van een VMBO eindexamen (niveau 2F) worden studenten geacht bekend te zijn met het **aflezen, interpreteren** en zelf **samenstellen** van grafieken. Studenten kunnen, onder andere, **data punten aflezen**, **conclusies trekken** gebaseerd op een grafiek en controleren of variabelen passen bij de **context** van een onderzoek. Onderwijs over misleidende grafieken ontbreekt echter nog waardoor het voor studenten moeilijk kan zijn om misleidende en correcte grafieken van elkaar te onderscheiden.

Veel gemaakte fouten:

- Gebruik van onjuiste grafiektype
- Niet alle datapunten tonen ('cherry picking')
- Y-as niet op nul laten beginnen
- Gebruik van verwarrende kleuren (tegen kleurnorm in gaan)
- 3D-effecten
- Omgekeerde Y-as
- Onjuiste schaal

Kernconcepten bij het interpreteren van grafieken:

Concepten in oplopende moeilijkheid:

- Data
- Context van een grafiek
- Verdeling
- Centrum
- Variatie of spreiding en variabiliteit
- Associatie
- Correlatie vs. causatie

Mogelijke werkvormen:

- Introductie les: introduceer de belangrijkste kernconcepten.
- Quiz met correcte en misleidende grafieken.
- Discussiegroep: welke invloed hebben misleidende grafieken op mij?
- Grafieken maken: studenten ontwerpen zelf (misleidende) grafieken en presenteren deze klassikaal.
- Misleidende grafieken zoeken: foto's maken van misleidende grafieken op sociale media en deze klassikaal bespreken.

Leerdoelen:

De studenten kunnen aan het einde van de les:

1. Verschillende soorten misleidende grafieken herkennen
2. Van verschillende soorten misleidende grafieken uitleggen waarom ze kunnen leiden tot misinterpretaties
3. Controleren of datavisualisaties in de media misleidend zijn. Ze denken kritisch na over datavisualisaties in de media en nemen niet zomaar alle informatie aan.

Extra informatie over (misleidende) grafieken:
Vijf studenten van Universiteit Leiden hebben samen met Universitair docent Sanne Willems een lesprogramma over misleidende statistieken ontwikkeld. Het lespakket bevat vier lessen die gecombineerd of los ingezet kunnen worden. Elke les bekijkt het onderwerp vanuit een andere hoek. Voor inspiratie of het downloaden van lesmateriaal, kijk op de volgende website:

- https://www.klascement.net/organisatie/32493/

Extra informatie over misleidende data visualisaties is te vinden op de volgende websites:

- https://nieuwscheckers.nl/grafieken-drie-soorten-misleiding/
- https://blog.vvsor.nl/2022/02/moet-je-altijd-bij-het-begin-beginnen
- https://ed.ted.com/lessons/how-to-spot-a-misleading-graph-lea-gaslowitz
- https://journals.sagepub.com/doi/10.1177/15291006211051956
- https://nos.nl/nieuwsuur/collectie/13903/video/2451256-de-misleidende-berichten-over-het-klimaat
- https://wisbase.nl/document/1656
- https://www.data-to-viz.com/

Extra informatie over kernconcepten bij het interpreteren van grafieken:

- https://www.nctm.org/uploadedFiles/Standards_and_Positions/Common_Core_State_Standards/Math_Standards.pdf of https://www.nctm.org/ccssm/

Extra informatie over hulpmiddelen bij het visualiseren van data:

- Boels, L. (2019). FIzier. Wat elke docent zou moeten weten over histogrammen, Euclides 94(4), 10–13. https://archief.vakbladeuclides.nl/bestanden/094_2018-19_04.pdf
- Boels, L (2022). Interpreteren van histogrammen en stippengrafieken, Euclides 97(5), p. 20. https://archief.vakbladeuclides.nl/jaargang_097.html

Ruimte voor notities:

....
....
....
....
....
....
....
....
....
....
....
....
....
....
....
....

Gemaakt door: Maaike Boele ten behoeve van afstudeeronderzoek voor de master Biologie en Wetenschapscommunicatie naar grafiek geletterdheid van MBO studenten.
Begeleid door: Lonneke Boels, Sanne Willems en Winnifred Wijnker

4. Interview protocol (in Dutch)

Vragen naar achtergrond docent:

- Hoeveel jaar geef je al rekenles?
- Welke opleiding heb je gevolgd?
- Heb je in het verleden wel eens les over misleidende grafieken gegeven?
- Maak je gebruik van een docentenhandleiding?
- Mag ik je leeftijd vragen?

Eventueel algemene vragen:

- Wat vond je van de les in het algemeen?
- Denk je dat het leerzaam was voor de studenten?
- Belang van onderwijs over misleidende grafieken?
- Op eerste oog lijkt dat alleen de eerste fout (fout op yas beter wordt herkend) heb je hier een mogelijke verklaring voor?

Vragen over docentenhandleiding:

- Heb je eerder gebruik gemaakt van een docentenhandleiding, zo ja hoe vaak en eventueel voor welke vakken?
    - ➢ Verschilt deze handleiding van eerder gebruikte handleidingen?
- Waarom denk je dat het hebben van een docenten handleiding wel of niet kan helpen met het voorbereiden van een les?
    - ➢ Is het dan minder werk/kost minder tijd voor je om zelf een les over misleidende data visualisaties te geven?
    - ➢ Is het genoeg informatie of heb je liever een syllabus met uitgebreider lesmateriaal en waarom?
- Als je dan kijkt naar de handleiding die ik heb gegeven aan het begin van de les: Welke aspecten vind je goed/handig
- Zijn er dingen die je mist op de handleiding die je zouden helpen met het voorbereiden van een les over misleidende data visualisaties?
- Zijn er aspecten op de handleiding die onduidelijk zijn en zo ja waarom welke?
- Heb je suggesties voor verbeteringen?