KU LEUVEN

FACULTY OF SCIENCE

Show Me Your Knowledge

Improving data recall by eliciting one's prior knowledge and comparing it with the correct data using storytelling techniques

Jonathan DELMEIREN

Supervisor:
Prof. dr. K. Verbert
Assessors:
Dr. D. Sileo
Prof. dr. A. Kimmig
Mentor:
D. Rojo Garcia

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Applied Informatics

Academic year 2020-2021

Preface

Throughout the writing of this thesis I received a great deal of support and assistance from a lot of people that I would like to sincerely thank. First and foremost I would like to express my gratitude to my mentor, Diego Rojo Garcia, for his invaluable expertise, flexibility and his guidance through each stage of the process. Next, I want to thank Prof. dr. Katrien Verbert for the opportunity to research in the interesting field of Information Visualisation and for the very much appreciated feedback after every presentation. Finally, I want to thank every participant of both user studies for their time and all my friends and family for supporting me during this challenging year.

Contents

Abstract

Visualisations that ask users to predict data before seeing it, are not a common occurrence. This is surprising, given that eliciting this prior knowledge increases the recall and comprehension of the data significantly [2]. Part of the reason might be the lack of research on how to adapt data visualisations to elicit the estimates of the user and compare them to the actual data. In this thesis we fill a part of this gap in the research literature, by focussing on this task for bar charts specifically. The first goal of this thesis is therefore to design a bar chart visualisation that is able to elicit user estimates and compare them to the correct data. The second is to evaluate whether this visualisation improves data recall.

A prototype is created by making adjustments and adding storytelling techniques to the best comparative layout for bar charts according to the state-of-the-art [6]. These storytelling techniques are added because we hypothesise that they further increase data recall by adding additional focus to the gap between user estimate and the correct data. We evaluate this prototype in a user study that consists of semi-structured interviews with five participants. After applying changes to the prototype based on the results of the first user study, we have a visualisation to use in the second user study that evaluates its effect on data recall.

In the second user study, we measure recall accuracy for three groups: a control group, who only sees the data in a regular bar chart, and two prediction groups where one of them has storytelling techniques added to the comparison visualisation and the other doesn't. The results show that data recall accuracy significantly improved when participants predicted the data and compared their estimates to the actual data. The addition of storytelling techniques did, however, not increase the data recall accuracy.

These results pave the way for a transition from static data visualisations to belief-driven ones in fields like digital media and education. To get the ball rolling, we developed an online tool to create these visualisations in a fast and easy manner, so authors and educators can start using them to increase data recall and comprehension for their readers and students.

Samenvatting

Het komt niet veel voor dat visualisaties gebruikers vragen om data te voorspellen voordat ze de correcte data te zien krijgen. Dit is verrassend, want het reflecteren op je huidige kennis over een onderwerp zorgt voor een sterke toename in het verwerken en kunnen onthouden van die informatie [2]. Een mogelijke reden hiervoor, zou het gebrek aan onderzoek kunnen zijn naar de mogelijke manieren om de huidige kennis van een gebruiker uit een welbepaalde visualisatie te halen. Verder moet er ook telkens een goede manier gevonden worden om de gebruiker te doen reflecteren over deze kennis, meestal door het vergelijken van zijn schattingen met de correcte data. Deze thesis wil een deel van dit gat in de huidige literatuur vullen door deze taak specifiek voor staafdiagrammen te onderzoeken. Het eerste doel van deze thesis is dus het ontwerpen en creren van een staafdiagram waarmee de gebruiker zijn schattingen van de data kan ingeven en kan vergelijken met de correcte data. Het tweede doel is om te onderzoeken of de gebruiker de data in kwestie al dan niet beter kan onthouden als hij gebruik maakt van de ontworpen visualisatie.

Er wordt een prototype ontwikkeld door enkele aanpassingen te maken en visuele verteltechnieken toe te voegen aan de beste vergelijkende staafdiagram visualisatie uit de bestaande literatuur [6]. We voegen deze verteltechnieken toe, omdat we hypothetiseren dat deze voor een extra verbetering in het onthouden van de data zouden kunnen zorgen, omdat ze zorgen voor een grotere focus op het gat tussen de voorspelling van de gebruiker en de correcte data. Het prototype wordt gevalueerd door middel van een gebruikers studie die vijf participanten interviewde over het design van de visualisatie. Hieruit kwamen enkele aanpassingen voor ons staafdiagram. Eens die waren toegepast, was onze visualisatie klaar voor gebruik in de tweede gebruikers studie. In deze studie werd de nauwkeurigheid van het zich herinneren van de gegevens gemeten voor drie groepen: de eerste is een controlegroep waarin de participanten de data simpelweg in een normaal staafdiagram bekeken, terwijl er bij de andere twee een schatting over de data moest worden gemaakt. Bij n van die twee schatting-groepen werd dan het verschil tussen de schatting en de echte data getoond aan de hand van visuele verteltechnieken en de andere niet. De resultaten toonden aan dat het onthouden van de data significant verbeterde als de gebruiker eerst de data schatte alvorens die schatting te vergelijken met de correcte data. Het gebruik van de verteltechnieken zorgde niet voor een significante verbetering.

Deze resultaten leggen de basis voor een transitie weg van statische informatie visualisaties naar de interactieve, schatting-gedreven visualisaties in velden als media en onderwijs. Om de bal aan het rollen te krijgen, ontwikkelden we een online tool waarmee auteurs en onderwijzers zelf dit soort visualisaties kunnen creren op een snelle en gemakkelijke manier, om zo hun lezers of leerlingen beter de informatie te kunnen laten onthouden en begrijpen.

List of Figures

List of Tables

Chapter 1

Introduction

Data visualisations have always played an important role in news media, but before the introduction of personal computers and the internet, they were static and couldn't be interacted with. Now that users' attention has become a coveted commodity online, the visualisations are being adapted to grab more of that attention. One of the main ways in which this is done is by adding storytelling techniques like for example animations, highlighting or user interaction.

A relatively new form of user interaction is letting a user predict the data before seeing it. One example is the 'You Draw It' visualisation by the New York Times [1] shown in Figure 1.1. In these so-called 'belief-driven visualisations' [7], the user is asked to predict the data by interacting with the visualisation, before comparing their estimate with the correct data. Multiple studies in cognitive psychology have shown that interacting with one's internal beliefs or prior knowledge of a dataset, is an important part in the interpretation process [8, 9, 10] and that expressing that prior knowledge can even improve a user's ability to adjust their current beliefs based on new information [11, 12]. Kim et al. [2] were the first to confirm these theories in relation to belief-driven visualisations. They found that expressing prior knowledge about the data by visually estimating it, significantly improved the recall and comprehension of that data compared to a control group. When combined with self-explanation or personalised feedback, this improvement increased even further.

These results, however, were only confirmed for slope graphs (Figure 1.2), opening the door for research on whether or not they are generalisable for other types of visualisations. Van Rooy [13] tested the same hypothesis for choropleths, a discrete geospatial visualisation, but found no significant improvement in data recall. Further research is needed to examine whether eliciting prior knowledge only positively affects specific visualisations like slope graphs or if discrete geospatial visualisations are an exception to the rule. In this thesis, we perform a user study similar to the one performed by Kim et al. [2] to test the hypothesis for another popular and more versatile visualisation type: bar charts. Concretely, we propose the following research question:

> RQ1: What is the effect of visually eliciting and reflecting on one's prior knowledge on data recall in bar charts?

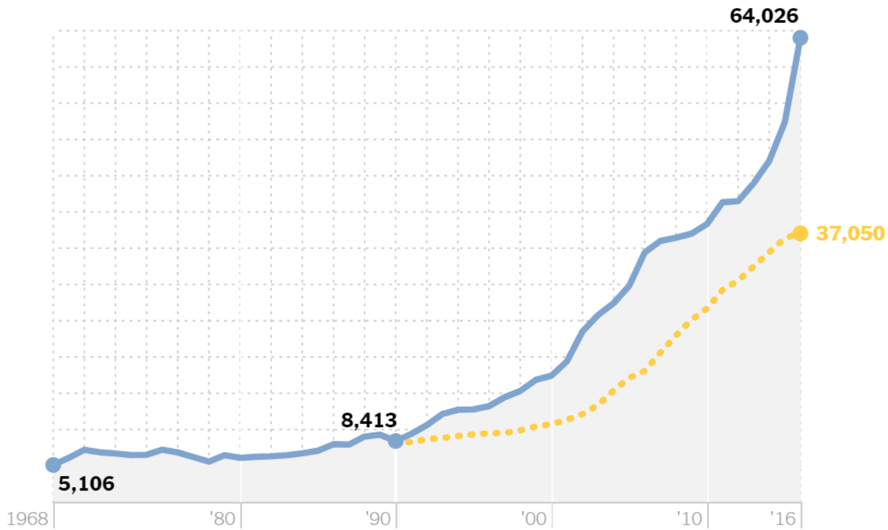

Figure 1.1: Belief-driven visualisation 'You Draw It' by the New York Times. The yellow dotted line represents the user's estimate while the blue line is the correct data. Image from [1].

Kim et al. [2] also researched the effect on data recall for two ways to focus the user's attention on the gap between their estimate and the correct data. They found that both self-reflecting and providing personalised feedback increased data recall compared to only estimating the data. When discussing future work they stated that: *"Animating feedback [...] may be particularly effective for drawing a user's attention to the gap."* [2]. Building upon this, we propose two sub-questions for the first research question:

> RQ1.1: What is the effect of visualising the gap between one's prior knowledge and the correct values on data recall?
>
> RQ1.2: What is the effect of visualising the gap between one's prior knowledge and the correct values using storytelling techniques on data recall?

Each of the two sub-questions evaluates a visualisation that elicits the user's prior knowledge and visualises the gap between estimate and correct data. However, the visualisation to answer RQ1.2 shows this gap by applying storytelling techniques to the comparison while the one for RQ1.1 shows it in a static manner.

Because the design space for creating belief-driven visualisations is relatively unexplored, creating a bar chart visualisation that elicits and reflects on prior knowledge is not trivial. When discussing future work, Kim et al. [2] proposed to drag the bars in a bar chart to set their height, but this elicitation method has not yet been tested. Regarding the comparing of estimates with correct data, research has been done on what the best comparison bar chart visualisation is to compare data series of equal importance, but none have answered this question for data series that are not of equal importance like the ones in our case. This is why we propose a second research question that we answer by conducting a user study that evaluates a prototype created based on current literature:

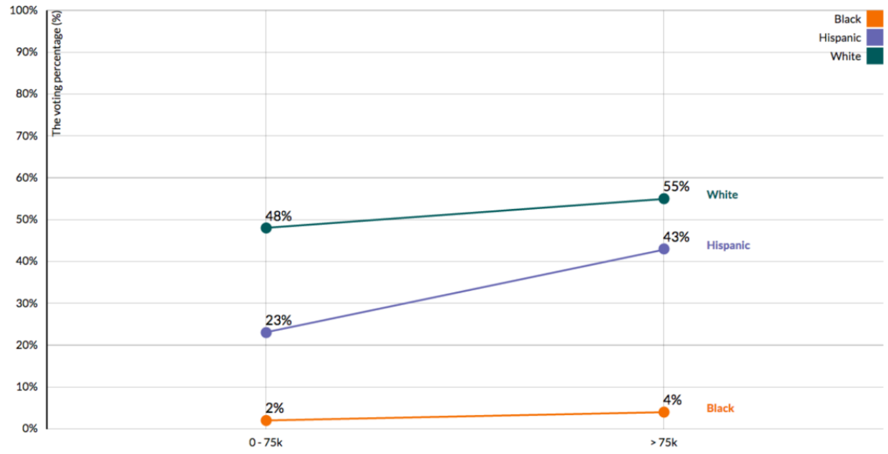

Figure 1.2: Example of a slope chart from the research by Kim et al. [2].

> RQ2: What bar chart visualisation is best for comparing the differences between user estimates and the correct data?

We expect to find that visually eliciting and reflecting on one's prior knowledge in bar charts significantly increases data recall based on the results in the paper of Kim et al. [2]. We also expect that applying storytelling techniques to the visualisation that compares a user's estimate with the correct data, has a greater positive effect on data recall than its static variant. If our expectations are met, the created visualisation could become a replacement for static bar charts in online media to create more interactive and memorable experiences for readers everywhere.

This thesis is structured as follows. Chapter 2 gives an overview of the related work that has been done on the topics of improving data recall by prediction, comparing data in bar charts and storytelling techniques. Based on these related works, we design and develop a prototype visualisation in Chapter 3. Chapter 4 discusses the selection of the dataset that is used in both user studies. The first user study is outlined in Chapter 5 and evaluates the prototype designed in Chapter 3. Based on the results, we apply changes to the prototype to create the comparison bar chart visualisation that serves as the answer to RQ2. The second user study makes use of this comparison bar chart visualisation to find an answer to RQ1 and its two sub-questions. How the study is conducted, together with an analysis of the results is detailed in Chapter 6. In Chapter 7 we interpret the results and their possible practical applications, together with the limitations of the thesis and potential future works. Finally, in Chapter 8, a conclusion to the thesis is drawn.

Chapter 2

Related Work

In this chapter, we examine prior works related to both research questions. In Section 2.1 we discusses the research that has been done on estimating data before seeing the actual data and self-explaining the gap between the two. In Section 2.2, the different ways of comparing data in bar charts are explored and we look at which of those ways is best for our use case according to current literature. Finally, in Section 2.3 we discuss different storytelling techniques and how to apply them to different visualisations.

2.1 Improving Data Recall by Eliciting Prior Knowledge by Predicting

We know from prior work that interacting with one's internal representations or prior knowledge of a dataset while looking at, or interacting with, a visualisation of that dataset, is an important part in the interpretation process [8, 9, 10]. When a person is asked to express those internal representations, it can even improve their ability to adjust their current beliefs based on new information [11, 12].

Two papers, "Explaining the Gap" by Kim et al. [2] and "Show Me Your Knowledge" by Van Rooy [13] examined whether eliciting and reflecting on one's prior knowledge while interacting with a visualisation has a positive impact on data recall and by extension data comprehension. The original paper, "Explaining the Gap" [2], found that *"participants who are prompted to reflect on their prior knowledge by predicting and self-explaining data outperform a control group in recall and comprehension"* [2]. They first created an interactive visualisation to elicit the user's prior knowledge and then used controlled experiments to evaluate the effect on data recall. In these experiments the participants first estimated the data before visually comparing their estimates with the actual data.

The visualisation created by Kim et al. [2] was a slope chart (Figure 1.2) designed specifically to support data with a particular structure that was beneficial for the study. When discussing future work, they asked the question whether their results would hold up for different visualisations and estimation methods. They created a list of possible prediction tasks for different encodings and possible interactions (Figure 2.1). For bar charts they propose to drag the bar to set its height.

Van Rooy [13] tested the same hypothesis for choropleths, a form of geospatial visualisations. They concluded that *"there was only a slight improvement in data recall accuracy*

Task	Detail Task	Manipulation Component	Encoding	Possible Interaction
Predict Continuous (Quantitative) Variable	Predict Data Value	Mark (bar)	Bar chart	Drag up a bar to set height
		Mark (line)	Line chart	Draw a line
		Mark Attribute (color of areas)	Map (choropleth)	Brush on color over an area
Predict Categorical (Nominal, Ordinal) Variable	Predict Categorical Membership	Mark Attribute (color of areas)	Area chart	Brush on color over an area
			Pie chart	Brush on color over a sector
Predict Data Structure and Model	Predict Correlation / Fit	Mark (line)	Scatter plot	Draw a regression line
			Line chart	Draw a line
	Predict Cluster		Scatter plot	Draw a contour
			Dendrogram	Drag and drop element to the cluster
	Predict Connectivity		Node-link diagram	Draw an edge
	Predict Confidence Interval		Box plot	Draw a line to mark confidence interval

Figure 2.1: Possible prediction tasks for different encodings and possible interactions. Image from [2].

when using prediction and self-explanation. No significant conclusion could be taken" [13]. They, however, found a significant difference in variance between groups but no clear explanation as to why this happened.

Based on the results of these two papers, it is not yet clear that eliciting and reflecting on one's prior knowledge while interacting with a visualisation has a positive impact on data recall for any type of visualisation. Therefore, in this thesis, we test the same hypothesis but specifically for bar charts. We also add another hypothesis where we theorize that using storytelling techniques when comparing between user estimates and the actual data, further increases this positive effect on data recall.

2.2 Comparing Data in Bar Charts

Bar charts display data using a bar for each category they want to represent data of. The height (or length, since bar charts can be plotted vertically or horizontally) of a bar indicates the value for that specific category. When there is more than one value per category, visually comparing those can be done in multiple different ways.

2.2.1 Design Space

L'Yi et al. [3] describe five types of comparative layouts for bar charts in their paper "Comparative Layouts Revisited: Design Space, Guidelines, and Future Directions" [3]. The first two are *chart-wise* (Figure 2.2, left) and *item-wise juxtaposition* (Figure 2.2, right). With chart-wise, two charts are placed side by side so they can be compared as a whole, while item-wise is meant to compare individual bars. For both of these juxtapositions, there are six ways to place either the charts or the bars next to each other. In particular: adjacent, stacked, mirrored, grid, diagonal, and free-form. The most prominent ones (adjacent, stacked and mirrored) are shown in Figure 2.2.

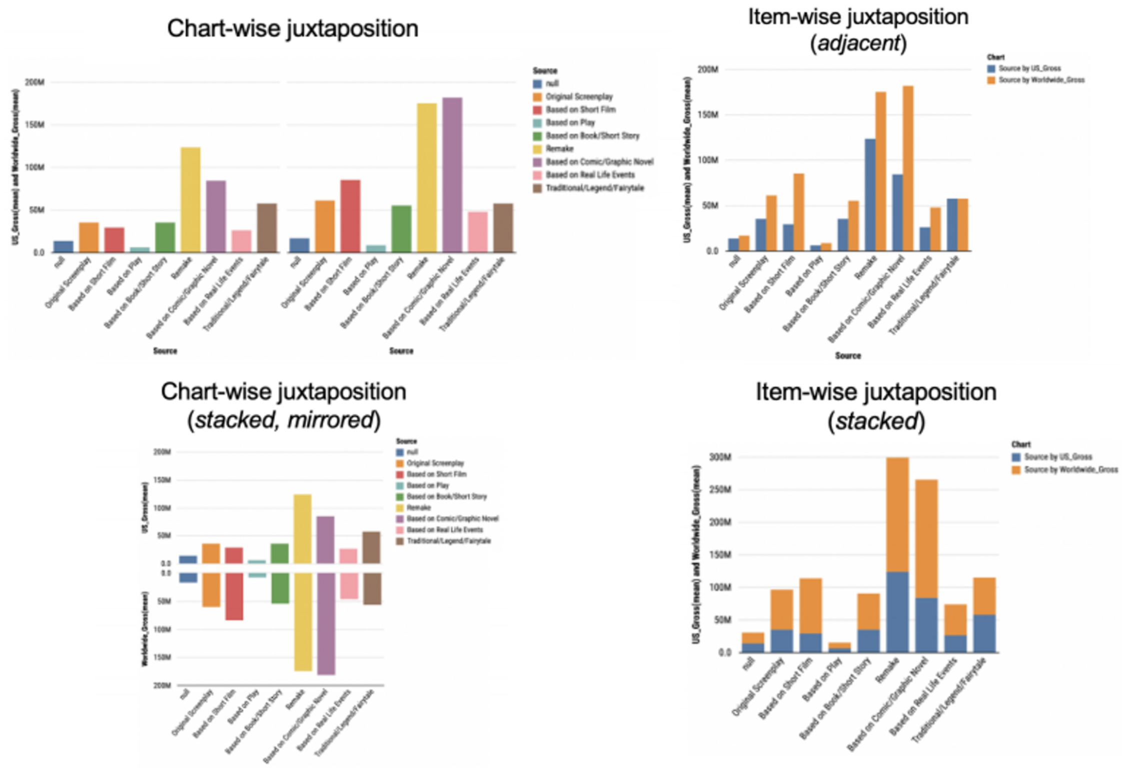

Figure 2.2: Four most prominent chart-wise (left) and item-wise (right) juxtaposition visualisations. Image from [3].

The next type of layout is *superposition* (Figure 2.3, left). It resembles item-wise juxtaposition in that it also combines all the data into one chart, but with the difference that bars can overlap. An example are bullet charts (Figure 2.3, right).

Explicit difference-encoding refers to a visual element that explicitly marks the relationships between two visualisations, for example the absolute difference between two bars. These elements can form a chart on themselves (Figure 2.4, left) or can be added to another visualisation (Figure 2.4, right). Explicit encoding is specifically designed to help with comparison tasks.

The *animated transition* category is the only visualisation that is not static. Instead, the design uses a temporal transition from one chart to another with the goal to make the differences stand out more.

The last category is *hybrid layout* and, as the name suggests, combines two or more of the other layouts to create a new layout. You could for example add explicit encoding to a grouped bar chart when the user should focus on comparing but still have the data of all the bars (Figure 2.4, right).

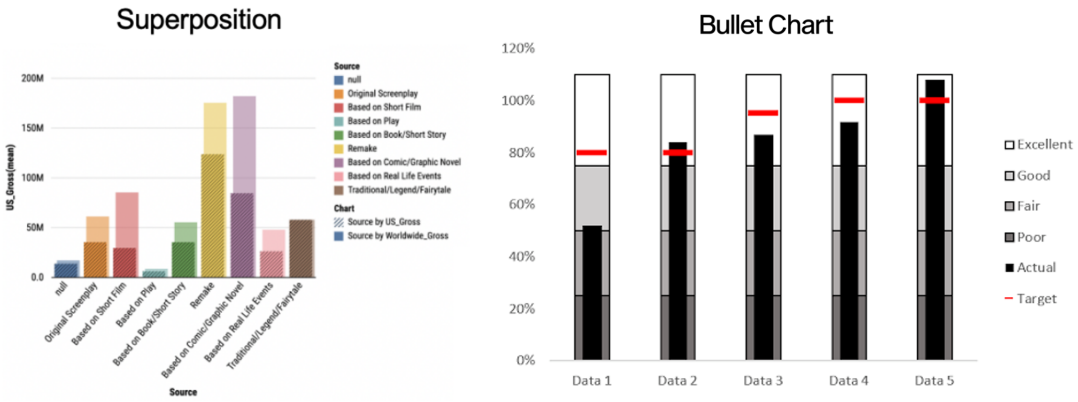

Figure 2.3: Superposition layout example from [3] and a bullet chart example from [4].

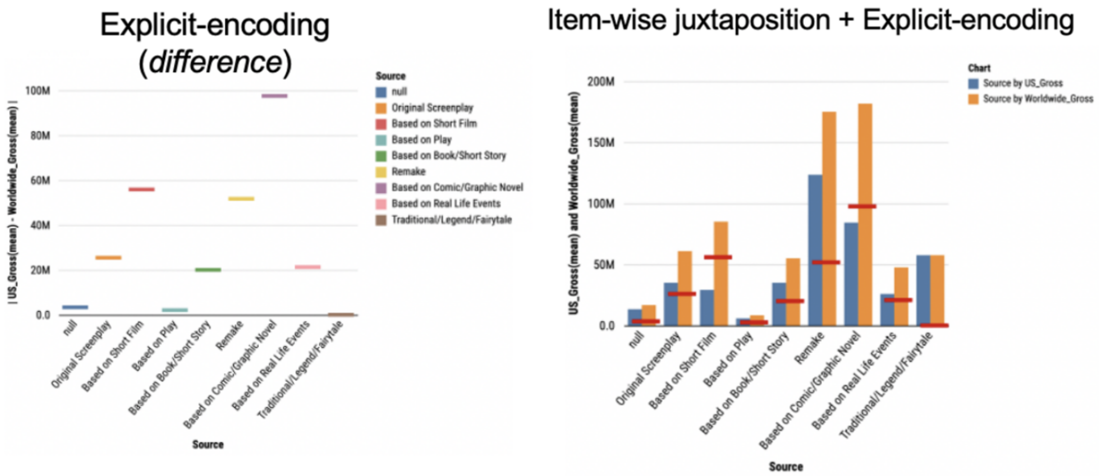

Figure 2.4: Examples of explicit difference-encoding. Images from [3].

2.2.2 State-of-the-art

We now look at three prior works that evaluated and compared multiple of these layouts for different comparison tasks. Carenini et al. concluded in their paper called "Highlighting Interventions and User Differences" [5] that adding explicit encoding to item-wise juxtaposition *"improves task performance as well as subjective user ratings"* [5]. They tested four different explicit encodings: bolding, de-emphasising, reference lines and connected arrows (Figure 2.5). All encodings except reference lines, improved the results on comparison tasks compared to the baseline. None of the other three clearly outperformed the others.

Where Carenini et al. [5] only examined static comparative layouts, Ondov et al. [14] also looked at animated transitions in their paper called "Face to Face: Evaluating Visual Comparison" [14]. They found that for comparison tasks in bar charts, these animated transitions outperformed any of the static comparative layouts. Of the four static comparative layouts they tested, superposition performed best, followed by mirrored stacked (chart-wise), adjacent (item-wise) and stacked (item-wise) bar charts.

In "What's the Difference?" by Srinivasan et al. [6] they evaluated four designs (Fig-

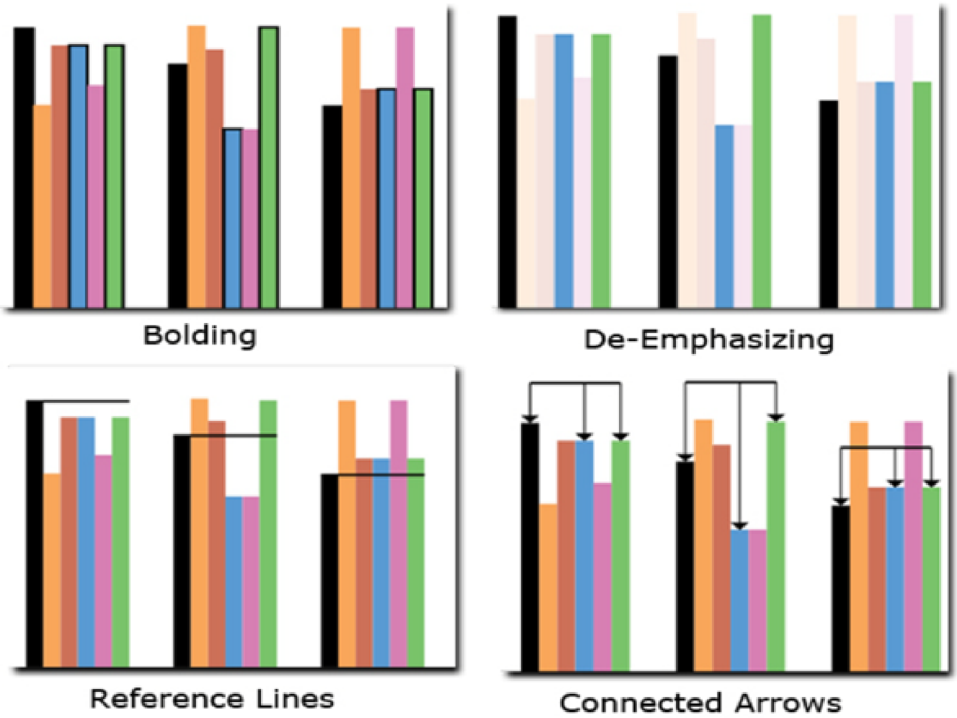

Figure 2.5: Four different explicit encoding tested by Carenini et al. [5].

ure 2.6) to compare two data series in a bar chart for multiple different comparison tasks. They determined that the grouped bar chart with difference overlay (Figure 2.6, c) and the single bar chart with difference overlay (Figure 2.6, d) performed equally well in comparison tasks but outperformed both singular chart types. Because the grouped bar chart with difference overlay (Figure 2.6, c) was the most liked by users, they concluded that this was the best visualisation for comparing data in bar charts.

To conclude, there is not one comparative layout for bar charts that always outperforms the others. Because Srinivasan et al. [6] evaluated designs for the same type of data that we use in this thesis, we use the visualisation that performed best in their study, a grouped bar chart with difference overlay as explicit encoding (Figure 2.6, c), as a baseline visualisation to create and compare our prototype with. This choice gets confirmed by the results of Carenini et al. [5] because it adds explicit encoding to item-wise juxtaposition and should therefore outperform a version without explicit encoding.

At the time of writing, there is no research on combining animated transitions with comparative layouts to satisfy the results found by Ondov et al. [14]. In the next section we look at research on how to include different storytelling techniques, like animations, to visualisations.

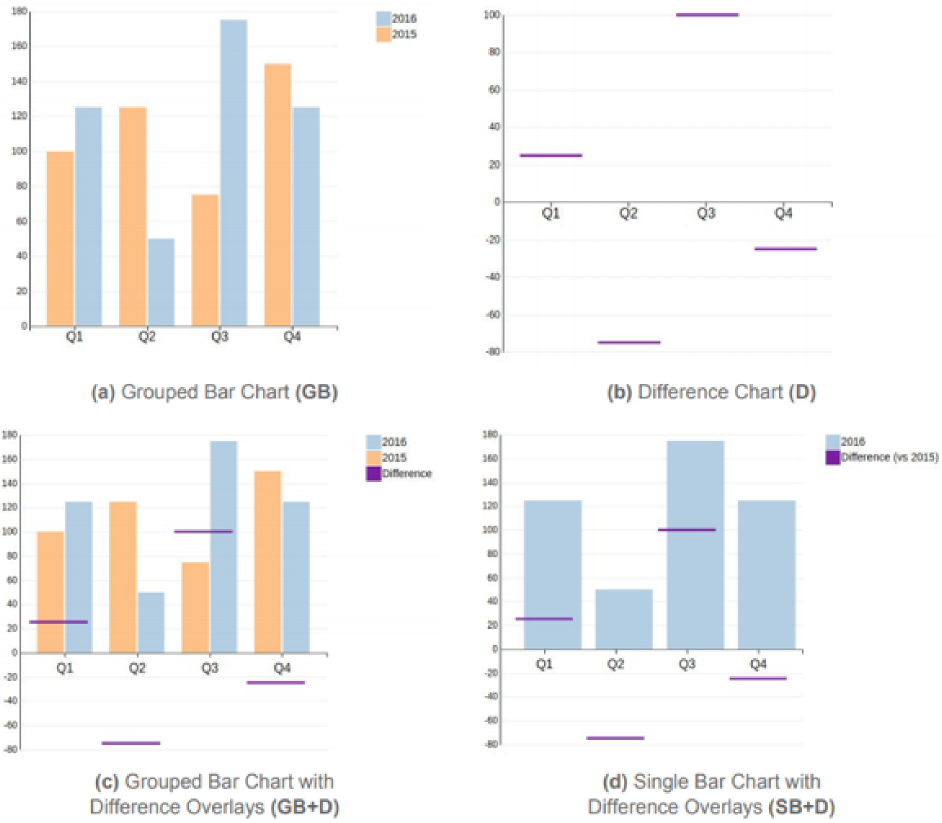

Figure 2.6: Four bar chart comparison visualisations evaluated by Srinivasan et al. [6].

2.3 Adding Storytelling Techniques to Visualisations

Segel et al. [15] formulated a design space around storytelling techniques in their paper called "Narrative Visualization: Telling Stories with Data" [15]. They drew on case studies from news media to visualisation research to identify distinct genres and applications of these narrative visualisations. For this thesis the visualisation genre is already determined as an annotated chart, which leaves only the following six design space elements.

The first one is *highlighting*, which is any visual mechanism that directs the user's attention to a specific element in the visualisation. The most used techniques are color, motion and framing.

Transition guidance is used when the visualisation needs to go from one static state to another. Transitions are almost always animations that bridge the gap between these states as smoothly as possible so as to not disorient the user.

The *ordering* of the elements in a visualisation is described as *"the ways of arranging the path viewers take through the visualisation"* [15]. This path can be decided by the author or at random through the order in which elements appear, disappear, are highlighted, etc. or it can be decided by the user by interacting with the visualisation.

Interacting with the visualisation is any way in which the user can manipulate or change the visualisation in combination with how the user is taught how to do that exactly. Examples are filtering, selecting, dragging and dropping, etc.

Anything related to textual communications towards the user falls under the category of *messaging*. These forms of text can range from short labels, annotations, headlines, etc. to actual texts like articles or introductions.

Finally, every narrative visualisation can be placed along a spectrum of *author-driven and reader-driven approaches*. On the side of the author-driven approach there is no interactivity, so the user has no control over the narrative at any point in time. For this approach, messaging is the most important aspect of the visualisation since the user can't explore the visualisation to find information. Situated on the other side of the spectrum, is the purely reader-driven approach. Here the author doesn't give the reader anything to go on and it's the task of the reader to explore the visualisation on their own. Interactivity is here the most important aspect of the visualisation.

Although there are countless possibilities on this spectrum, there are two hybrid models that have become popular. The first one is the *Martini Glass structure* which begins on the author-driven side of the spectrum but later opens up, just like the shape of a Martini glass, to a reader-driven stage. In this stage the user is free to explore the visualisation, because the author's intended narrative is already complete. The second model is called the *Interactive Slideshow*. Here the user gets to interact with the visualisation after every 'slide' of author-driven narrative.

Chapter 3

Prototype Design and Development

In this chapter we describe the design and development of our comparison visualisation prototype (Figure ?) that we evaluated in the first user study in Chapter 5. It is based on the baseline visualisation (Figure 2.6, c) that we describe in Section 2.2 and is the best comparative layout for bar charts according to the state-of-the-art [6].

In Section 3.1 we describe the different design choices we made to create the estimation part of the visualisation. In Section 3.2 we explain why the baseline comparison visualisation from Srinivasan et al. [6] was not the ideal visualisation for this thesis and what changes we made to get to our static comparison prototype. In Section 3.3 we provide two reasons for adding storytelling techniques to our comparison visualisation and describe what specific storytelling techniques were added. In Section 3.4 we describe the development of the prototype and Section 3.5 summarises how we got to our final prototype.

3.1 Estimation Bar Chart Prototype Design

Before being able to compare the user's estimates with the correct data, those estimates need to be elicited. Kim et al. [2] created a list of possible interactions to elicit user estimates from different prediction tasks for different encodings (Figure 2.1). For bar charts they proposed to drag the bar to set its height. We also found this to be the most natural way, so in our prototype a user could drag the top of a bar to its desired height to input their estimate for that bar. In the following subsections, we describe three additional design choices that had to be made to get the prototype estimation chart (Figure 3.1).

Starting Height

The first additional design choice was the starting height of the bars. In the prototype these were set to 10% of the maximum value on the y-axis. This was low enough as to not confuse the user that these bars were already at their correct height, but left enough room to grab and drag the top of the bar.

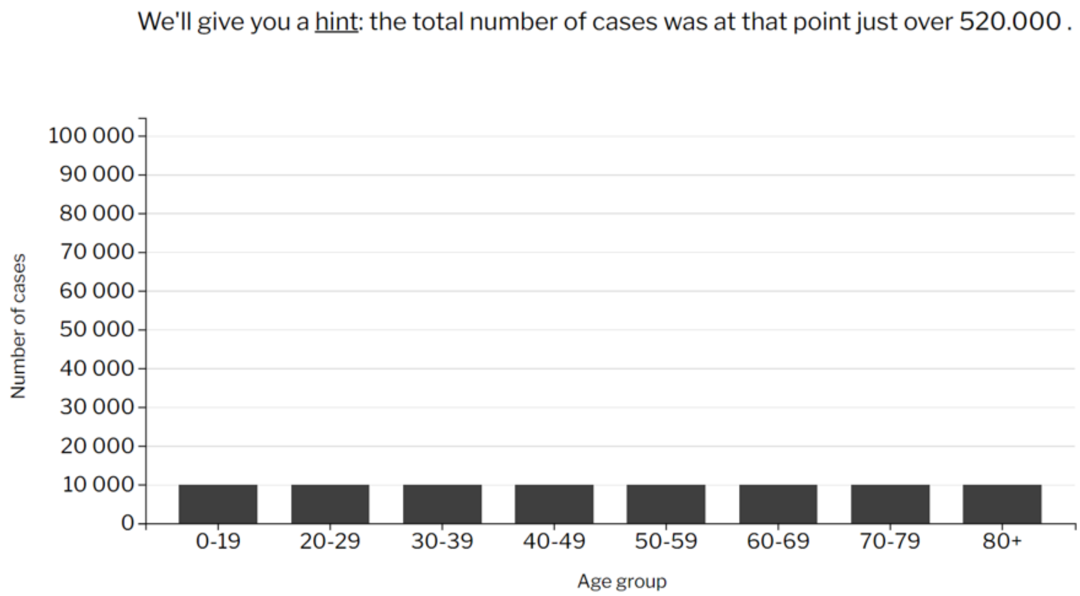

Figure 3.1: The estimation prototype chart. Users can input their estimate by dragging the top of each bar to its desired height.

Prediction Hints

Two other design choices fell into the category of what Kim et al. [2] call *"the contextualization mechanisms (how does the interface provide clues to constrain the user's prediction?)"* [2]. The first one is *prediction hints*. These provide direct clues to how or where a user should make a prediction. Hints can be textual (e.g. providing additional information about the data) or visual (e.g. giving part of the correct data already filled in in the visualisation). We adopted the first option and gave the user the sum of all correct values as a hint in a sentence above the visualisation.

Y-axis Scale

The second contextualization mechanism was the scale of the y-axis. Kim et al. [2] observed that:

> *"Users' predictions were quite sensitive to the axis range. When we presented the full 0-100% percentage range for percentage variables (e.g., the percentage of the U.S national budget of health care), users' estimates showed a bias toward the center in the plotting range. This effect was lessened when we trimmed the axis range based on the maximum value of the dataset, suggesting that users implicitly view the axis range as a clue to the data scale."* [2]

For this reason we kept the maximum y-axis value close to the correct value of the highest bar. The maximum y-axis value in the prototype was calculated as the highest correct value increased by one sixth of that same value. If, for example, the highest correct value was 120, the maximum y-axis value would have been 140.

Figure 3.1 shows the complete prototype estimation chart the user sees when asked to estimate the correct values of the bars.

3.2 Static Comparison Bar Chart Prototype Design

In this section we describe the design of the visualisation used to compare the user estimates with the correct values (Figure 3.2). The goal of this visualisation is to show the user the gap between their estimate and the correct data, while making sure that they remember the correct data as well as possible.

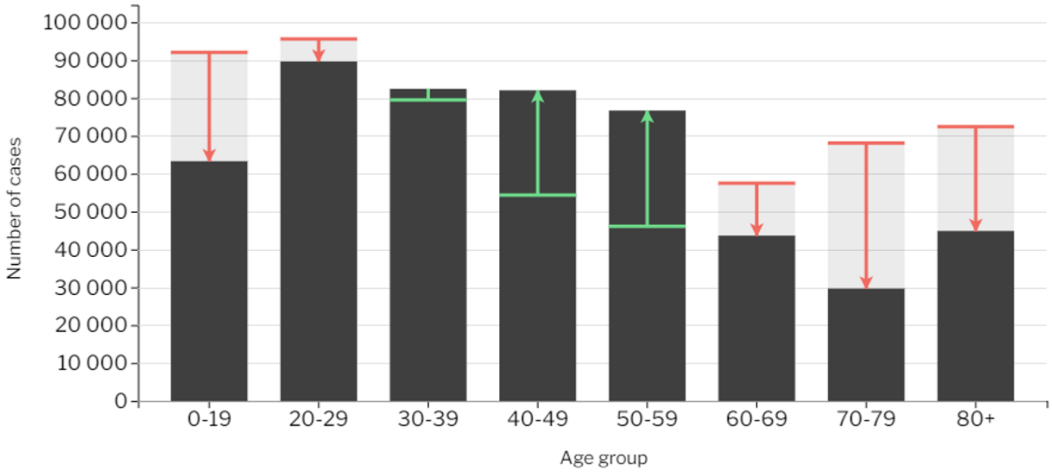

Figure 3.2: The comparison prototype chart.

In Section 2.2 we determined that, based on the work of Svirissan et al. [6], a grouped bar chart with difference overlay as explicit encoding (Figure 3.3) is the best bar chart visualisation for comparing two data series of equal importance. However, a key difference with the data in this thesis is that both data series are not of equal importance. The first series is the estimates of the user while the second is the actual, correct data. The estimates are only important to show the gap between the user's prior knowledge and the correct data, users should not remember these data points since they are incorrect. This means that placing them next to each other as equals would take attention away from the correct data, which would not be beneficial for data recall.

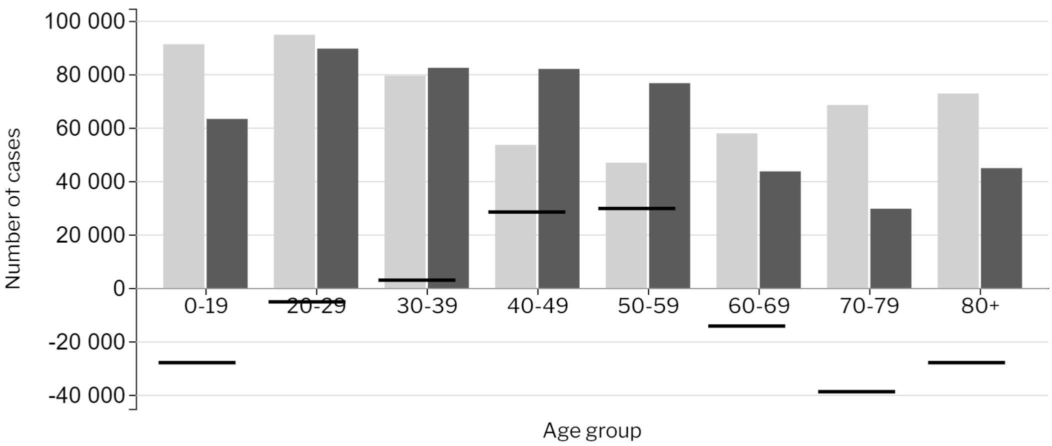

Figure 3.3: Baseline visualisation: a grouped bar chart with difference overlay as explicit encoding.

For this reason we proposed three changes to the baseline comparison visualisation (Figure 3.3) to create our static comparison prototype (Figure 3.2) that we evaluated in the first user study. The first change was to remove the bars of the user estimates and only keep the bars of the correct data. This way the bars of the data that we wanted the user to remember, namely those of the correct data, could not be confused with those of the estimates.

The second change was to explicitly encode the user's estimates as lines on the bars instead of the difference between their guess and the correct data. This more clearly shows the actual 'gap' between the two and makes more sense in an accompanying animation.

The third and final change was the addition of an arrow from the line of the user estimate to the top of the bar. This increased the readability of the chart. The direction of the arrow indicates whether it was an over or underestimation. In addition to direction, we also color coded the arrows to display this. Upwards and green meant an underestimation and downwards and red an overestimation. We used green and red, even though they are not colorblind-safe, because they are very well known encodings for upwards and downwards.

After applying these three changes we get our static comparison bar chart prototype as shown in Figure 3.2.

3.3 Adding Storytelling Techniques to the Static Comparison Prototype

The goal of adding storytelling techniques to the visualisation is twofold. First, we hypothesise that it might draw more of the user's attention to the gap between their estimate and the actual data. This was also hypothesised by Kim et al. [2] when discussing future work:

> *"Animating feedback, such as by dynamically moving marks added by the user to their true positions, or adding textual feedback to prediction errors point by point, may be particularly effective for drawing a user's attention to the gap."* [2]

Our second hypothesis based on prior research [16, 17] is that it might increase the user's engagement with the visualisation which in turn would have a positive effect on the ability to recall the data.

Before adding concrete storytelling techniques to the prototype, we describe the general idea. We wanted to tell a story where we started from a 'filled in' estimate chart as shown in Figure 3.4, where the user had estimated the values for the bars by dragging them to their desired height. From this estimation chart, we wanted to get to the static comparison bar chart shown in Figure 3.2. During this story we wanted to draw the user's attention to each of the gaps between their estimates and the correct data, while ensuring that, in the end, they had all the information needed to remember the correct data. Because we wanted the user to focus on each of the gaps separately, the bars needed to be animated or focused on in a bar-per-bar fashion. We now go over the design space as described in Section 2.3 to outline how we applied the different storytelling techniques.

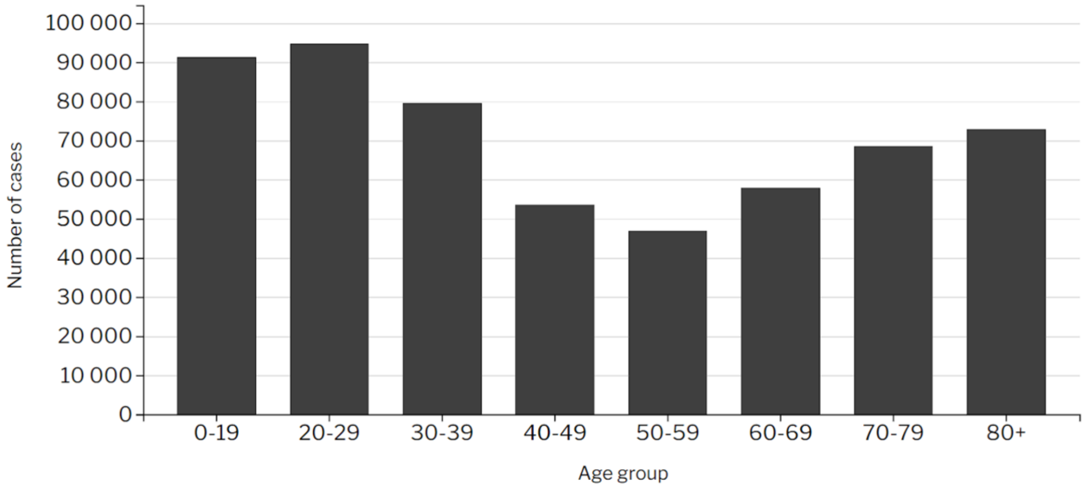

Figure 3.4: An estimation prototype chart after a user estimated the data by dragging the bars to their desired height.

Transition Guidance

The most essential technique is transition guidance, which we used twice in our visualisation. First, it refers to the animations that are used to go from the user estimate to the correct value and second when we go from one bar to the next. Regarding the former, we had to go from a bar like Figure 3.5 (a) to one like 3.5 (d). The most logical and intuitive way to achieve this was to grow (or shrink) the bar, leaving a line at the estimated value and drawing an arrow that shrinks (or grows) together with the bar (Figure 3.5 (a) to (d)). The latter is a transition between bars, but this is less of a transition and more of a change in highlighting.

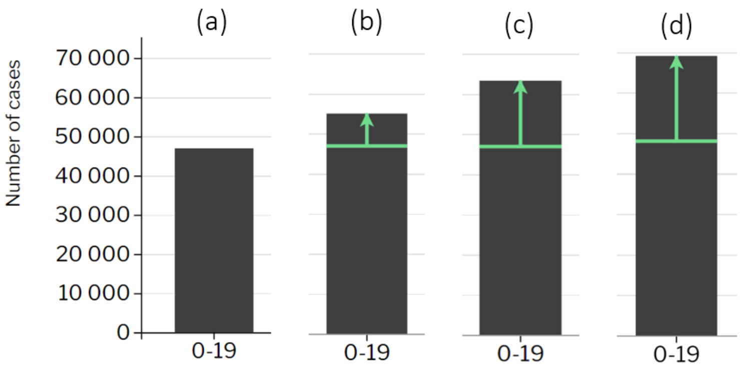

Figure 3.5: The growing of a bar from the user's estimate to the correct value

Highlighting

At any given time during the full animation, there is only one bar highlighted: the bar that is currently undergoing its animated transition. When a bar is highlighted it means that the other bars are greyed out as shown in Figure 3.6. After the animation has finished, the user can hover over a bar to highlight it again.

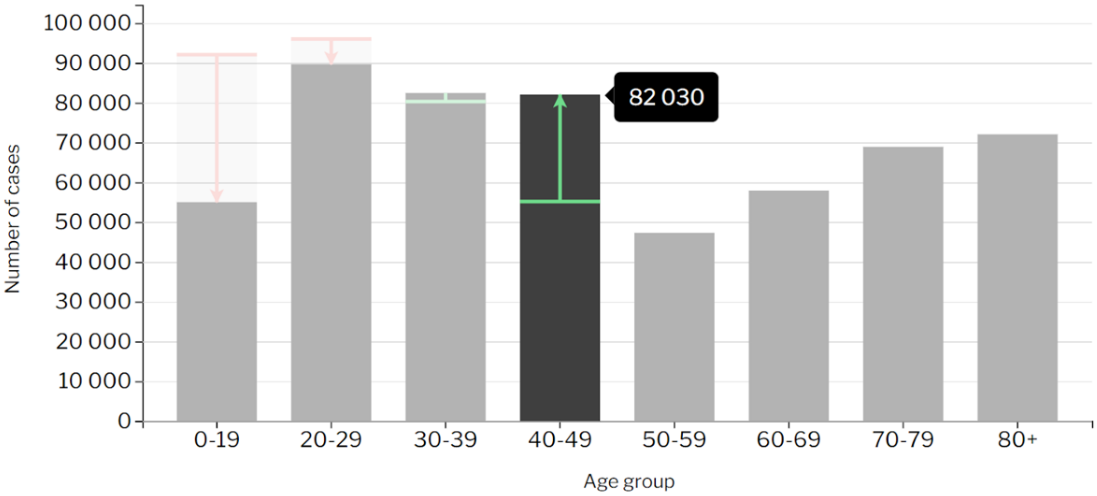

Figure 3.6: Comparison visualisation during the animation. Currently, the bar of the 40-49 age group is transitioning and is therefore highlighted.

Ordering

In Figure 3.6, the bars to the left of the highlighted bar have already transitioned while the ones on the right still need to do so. This indicates the left-to-right ordering of the narrative structure. This is the simplest, most obvious and most expected order for our target audience, who use the left-to-right writing system.

Messaging

We included two different messaging elements in the visualisation. The first is a tooltip next to a highlighted bar with the value that corresponds to the current height of the bar (Figure 3.6). During the animation of a bar, this value starts as the estimate of the user and ends on the correct value. During the growing or shrinking of the bar, the value changes together with the changing height, trying to display all values that are in between.

The second message is a sentence of textual feedback that appears underneath the chart on how the user estimated when they hover over a bar after the full animation has finished. The sentence reads for example: *"The correct value for 50-59 was 50 779 cases higher than you estimated."* (Figure 3.7).

Interactivity

Users can interact with the visualisation by hovering over the bars. This is only possible when all bars have transitioned. When a user hovers over a bar, it highlights that bar and displays the relevant textual feedback underneath the chart (Figure 3.7).

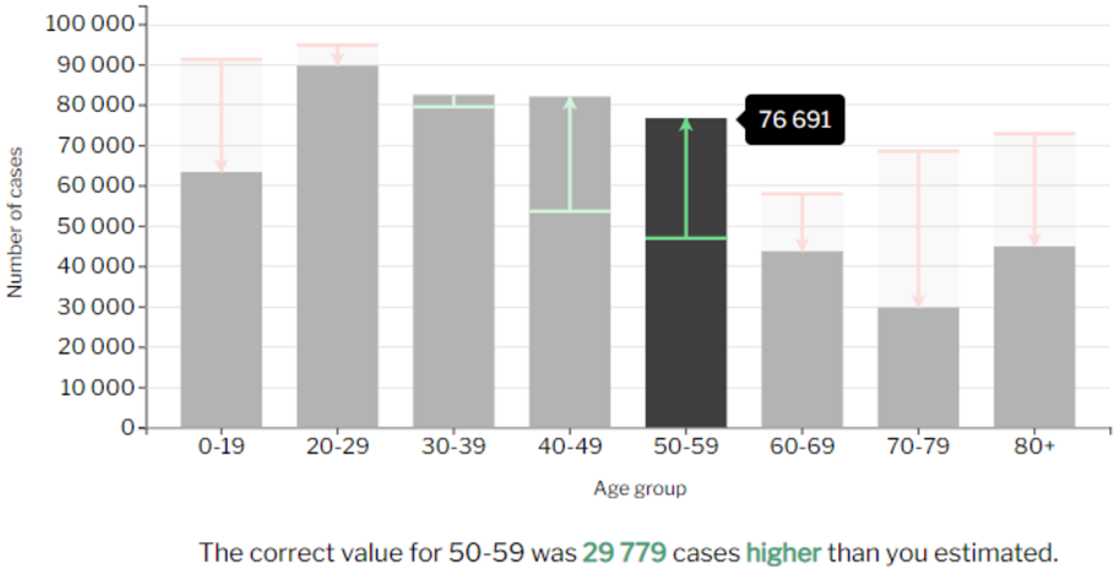

Figure 3.7: Comparison visualisation when the user hovers over a bar.

Author-Driven versus User-Driven Approach

We followed the Martini Glass approach and animated all of the bars immediately after each other without letting the user interact with the visualisation in between. After all the bars have transitioned, the user can freely interact with the visualisation to revisit any of the values by hovering over them.

3.4 Development

The visualisation was created in the programming language JavaScript with the help of the D3.js library[1]. The website around the visualisation was built with HTML and CSS and was hosted on Github Pages[2] for both user studies. The source code for the final visualisation can be found on the Github page with all supplementary materials for this thesis[3].

3.5 Complete First Prototype

When we put together all design choices from the previous sections, we get the complete first prototype in Figures 3.8 to 3.12. Figure 3.8 shows the initial chart where the user inputs their estimates by dragging the top of the bars to their desired height. When finished, it looks something like Figure 3.9. When the user decides they want to see the results, the animation starts by highlighting the leftmost bar and, in this example, shrinks the bar to the height of the correct value (Figure 3.10). When the transition of the first bar has finished, the second bar gets highlighted and the animation starts anew. When all the bars have transitioned, the highlighting drops away and we get the static comparison

[1]https://d3js.org/
[2]https://pages.github.com/
[3]https://github.com/JonathanDelm/ShowMeYourKnowledge

bar chart as shown in Figure 3.11. The user is now free to hover over any of the bars to highlight them again and reveal a sentence of textual feedback underneath the chart like in Figure 3.12.

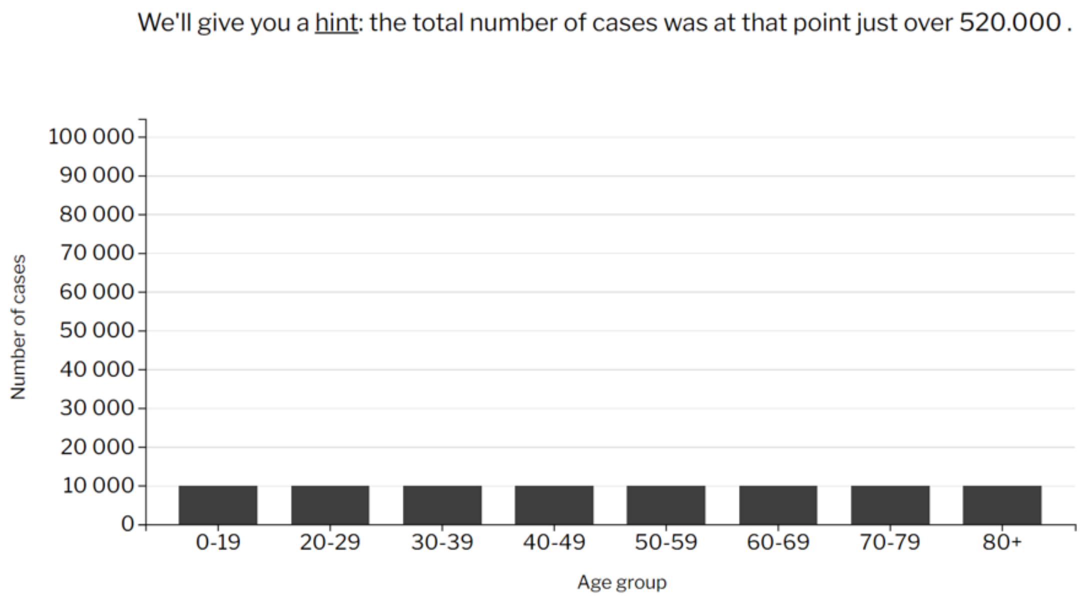

Figure 3.8: The initial chart where the user inputs their estimates by dragging the top of the bars to their desired height.

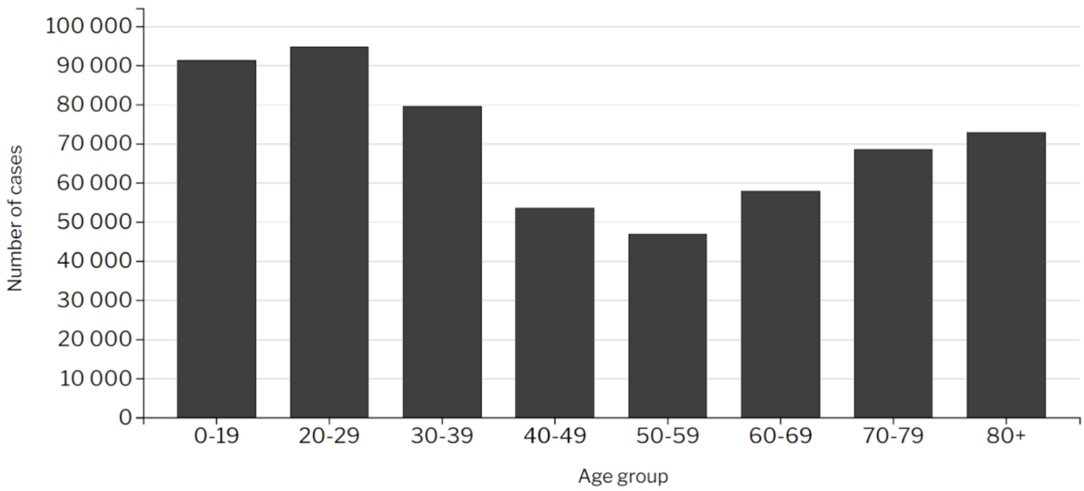

Figure 3.9: The estimation chart after a user estimated the data by dragging the bars to their desired height.

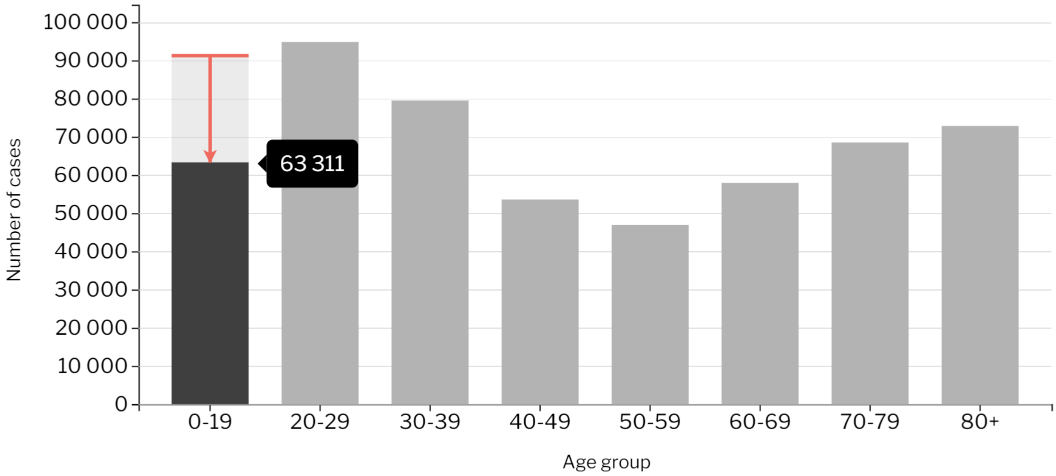

Figure 3.10: The animation starts by highlighting the leftmost bar and, in this example, shrinks the bar to the height of the correct value.

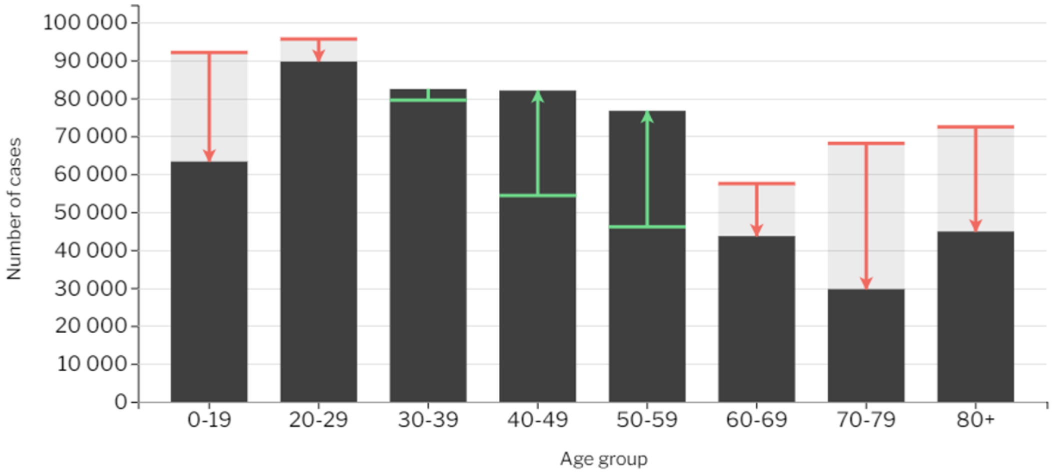

Figure 3.11: After the animation, we see the full comparison prototype chart.

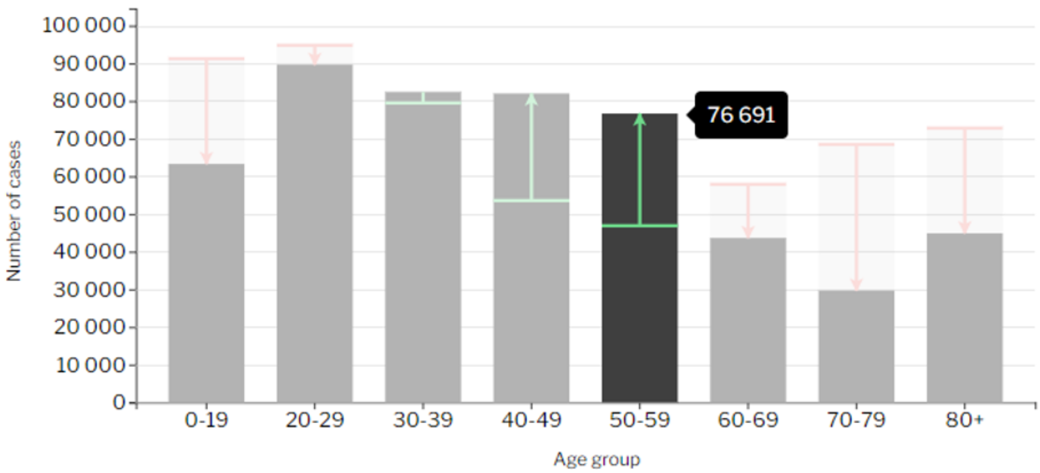

Figure 3.12: Comparison visualisation when the user hovers over a bar.

Chapter 4

Dataset Selection

In order to evaluate our prototype, we needed to select an appropriate sample dataset to use in both user studies. Kim et al. [2] found through a replication study that high familiarity datasets did not result in an increase in data recall. This highlights the importance of selecting the right sample dataset that is not too familiar to the target audience. In this chapter, we describe the process followed to pick a dataset.

In section 4.1, we describe our target audience. In section 4.2, we list the requirements that were used to evaluate different possible datasets. These requirements are only used to select a dataset to use in the two user studies in this thesis and are not requirements for datasets in practical applications. Finally, in section 4.3, we use these requirements to select a dataset. The appropriateness of this dataset was confirmed by a small evaluation in the first user study.

4.1 Target Audience

The designed visualisation is suggested as a static bar chart replacement in online media to improve data recall. Thus, the target audience is quite broad as it includes any person who reads articles online. However, since the participant recruitment process was planned to target Belgian residents, we focused on datasets that are relevant for them.

4.2 Dataset Requirements

We define ten desired requirements that we use to select a dataset for our user studies. The full list can be found in Appendix A together with all considered datasets. We describe three main requirements in more detail in the following subsections.

4.2.1 Relevance for Target Audience

To evaluate the effect of eliciting prior knowledge on data recall, the target audience needs to have some prior knowledge about the dataset. The data also needs to be interesting so they want to explore and interact with the visualisation. If we were to, for example, ask a group of Belgian married couples about the most popular Chinese dating apps, they probably wouldn't know anything about the data nor care about it.

4.2.2 Seven to Nine Data Points

In a 1956 paper called "The magical number seven, plus or minus two" [18] , Miller, G. A. concluded that the number of objects an average human can hold in short-term memory is 7 ± 2. Van Rooy [13] hypothesised that *"the [low] amount of bins could be one of the causing factors for the non-significance of the Data Recall experiment"* [13]. In his experiment, 20 out of 114 participants recalled all data points perfectly. He speculated that increasing the amount of bins, therefore making the task of recalling more difficult, might have made the positive effect of estimating more pronounced. To make the task of recalling in our user studies sufficiently difficult, but not too difficult, we required datasets to have a minimum of seven and a maximum of nine data points, at the high end of the 7 ± 2 spectrum.

4.2.3 Ordinal Data

If we were to select a dataset with nominal data, we would need to manually select an order. Picking this order adds an extra variable to our research that might influence the analysis and results of the study. To avoid this possible influence, we opt for ordinal data so the order of the bars is predefined.

4.3 Dataset Selection

The dataset we chose is the distribution of confirmed coronavirus cases in Belgium by age on Statista[1]. It fulfilled all ten requirements while also being an incredibly relevant topic at the time of conducting the user studies. For the first user study we used the data up to November 11th 2020 (Figure 4.1) and for the second up to May 1st 2021 (Figure 4.2).

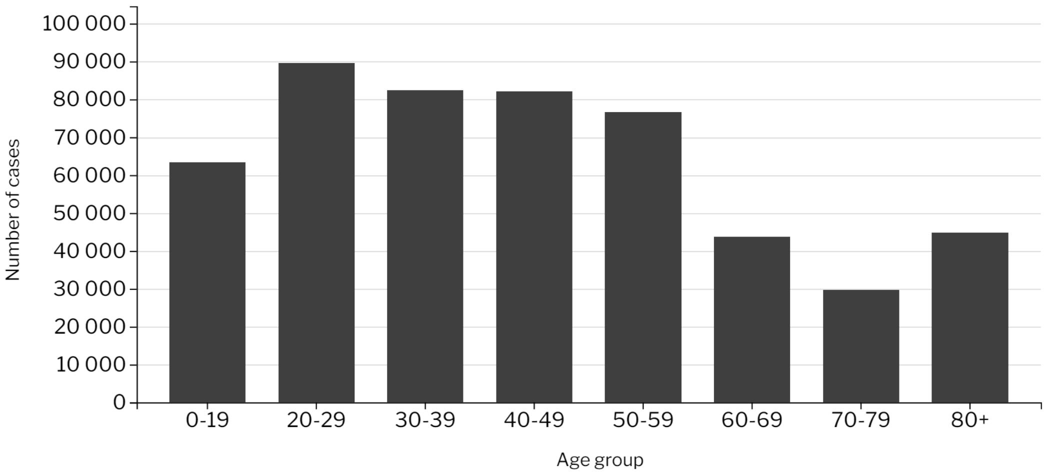

Figure 4.1: The distribution of confirmed coronavirus cases in Belgium by age up to November 11th 2020. Data from [1]

[1]https://www.statista.com/statistics/1114426/confirmed-coronavirus-cases-in-belgium-by-age/

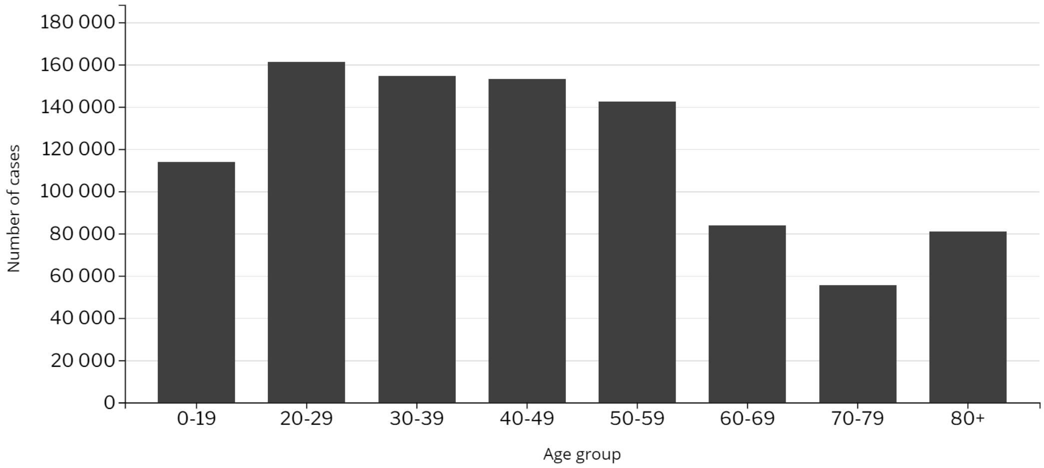

Figure 4.2: The distribution of confirmed coronavirus cases in Belgium by age up to May 1st 2021. Data from [1]

Chapter 5

First User Study

This chapter outlines the first user study. In this study, we evaluate the comparison visualisation prototype characterized in Chapter 3 (Figures 3.8 to 3.12). In Section 5.1 we describe the goal of the study. In Sections 5.2 and 5.3 we outline the methodology of the study and the concrete protocol that was used respectively. The results of the user study are discussed in Section 5.4. The final section, Section 5.5, summarises all the changes made to the prototype and shows the full visualisation that is used in the second user study.

5.1 Goal

The first user study aims at answering RQ2: *"What bar chart visualisation is best for comparing the differences between user estimates and the correct data?"*. For this, we devised certain changes to some of the design elements of the best comparative layout for bar charts according to the state-of-the-art [6]. The resulting visualisation is our comparison visualisation prototype (Figures 3.8 to 3.12). This user study aims at evaluating the applied changes in the context of comparing estimates with actual data.

Concretely, We want to receive feedback on the different design areas of our prototype. A list of all these design areas is provided in Table B.1 in Appendix B, together with their specific design elements, what questions we asked the user and, if relevant, some different design options that were considered during development.

5.2 Methodology and Setup

All interviews were conducted in a secure Microsoft Teams session where the participants could join without having to provide identification. Every interview started with an introduction where the goal of the study and practicalities were discussed. One of these practicalities was sending a GDPR information sheet (Appendix C) with all necessary information regarding their privacy. After the introduction, participants were asked five demographic questions: age, gender, highest earned degree, amount of contact with visualisations like bar charts and amount of online articles they read. Afterwards, they were sent a link to the website where the visualisation was hosted. The interview was conducted following the procedure described in Section 5.3. The introduction for the user study interview in its fully written out form can be found in Appendix D.

5.3 Interview Procedure

The interview started with a think aloud procedure to get user feedback without having to provoke it with specific questions. The user explored the visualisation and was asked beforehand to talk about everything they did and thought without being interrupted by the interviewer. This was done so the user could go at least once through the whole process like they would in a normal use case.

After that first exploration, the interviewer brought up any remarks made by the user during the think aloud procedure and asked them to go into more detail. The interviewer first tried to find explanations for remarks by asking questions like "Why do you think that?" or "What did you expect to happen?". With the goal to find things that needed fixing or were not clear to the user. Secondly, we wanted the user to provide concrete suggestions for improvements when giving negative remarks. So whenever the user made a negative remark, we wanted to find out whether they also had a suggestion that, according to them, would solve their problem. Therefore, every negative remark was followed up with: "And how would you solve this?" or "How would you change this?".

Next, we wanted to provoke input for other elements in the same design area according to Table B.1. When all the unprovoked user input about a certain design area was handled, but there were still some parts of that area that needed to be addressed, the interviewer introduced them with questions like: "And what do you think about this?" or "Since you said that you didn't like that, what is your opinion on this then?".

The next types of questions were meant to get comparative user input. Here the interviewer made the user compare different implementations when needed and/or possible. If, after all the previous questions, it was clear that the user disliked a certain implementation, but hadn't suggested all (or any) of the alternative implementations mentioned in Table B.1, the interviewer proposed the different implementations and asked if these were any better and if so, why. An example could be: "Here is another implementation of this element we considered. Do you think this is a better implementation? How so?".

Finally, we wanted to provoke input on unmentioned areas. If a user didn't give unprovoked input on a whole area, the interviewer introduced it to them manually.

5.4 Results and Changes to the Prototype

We interviewed five participants, three male (M1, M2, M3) and two female (F1, F2) with a mean age of 33.8 (SD = 15.79). The answers to the other three demographic questions are listed per participant in Table 5.1. On average, each interview took about one hour to complete. The results of the study are based on notes that the interviewer took during the interviews. The list of notes per design element per participant is provided on the Github page with all supplementary materials for this thesis[1].

In this section we go into detail about design elements where we made changes based on the results of the user study. In each of these subsections, we first discuss the results and afterwards reveal the changes we made to the visualisation based on those results. The remarks about the design elements that weren't changed are combined into one subsection where they are each discussed briefly.

[1]https://github.com/JonathanDelm/ShowMeYourKnowledge

Participant	Age	Highest degree	Contact with visualisation	Amount of online articles
M1	23	Bachelor	3 times per day	9 per day
M2	53	High School	2 times per week	10 per week
M3	23	Master	3-5 times per week	3-5 per week
F1	21	High School	1-2 times per day	1-2 per day
F2	49	Master	Once per day	5 per day

Table 5.1: Demographic information of the five participants of the first user study.

5.4.1 Prediction Hint

The first design element is the hint we gave participants to give them some perspective during the estimation process. For this we provided the following line of text above the chart: *"We'll give you a hint: the total number of cases was at that point just over 520 000"* (Figure 3.8). All five participants were glad they got a hint to help them, but all of them also said it was difficult to use the hint without some way of knowing what their current total was, to check whether they were close to the 520 000 or not. The other options were also generally liked without there being a clear favorite.

Because the hint is only meant to give perspective and be a guideline rather than a requirement that the user needs to fulfil, we remove it from the visualisation. As a replacement we now give the correct height of one of the bars when estimating (Figure 5.1). This way, a user can base their other guesses on this one bar, without having to start counting or changing some of the bars because they aren't at the correct total yet. Concretely, we decided on the bar for the 60-69 age group. This one is somewhat in the middle both in terms of age groups and number of cases and is also not one of the potentially 'surprising' results like for example the 20-29 group having the highest number of cases.

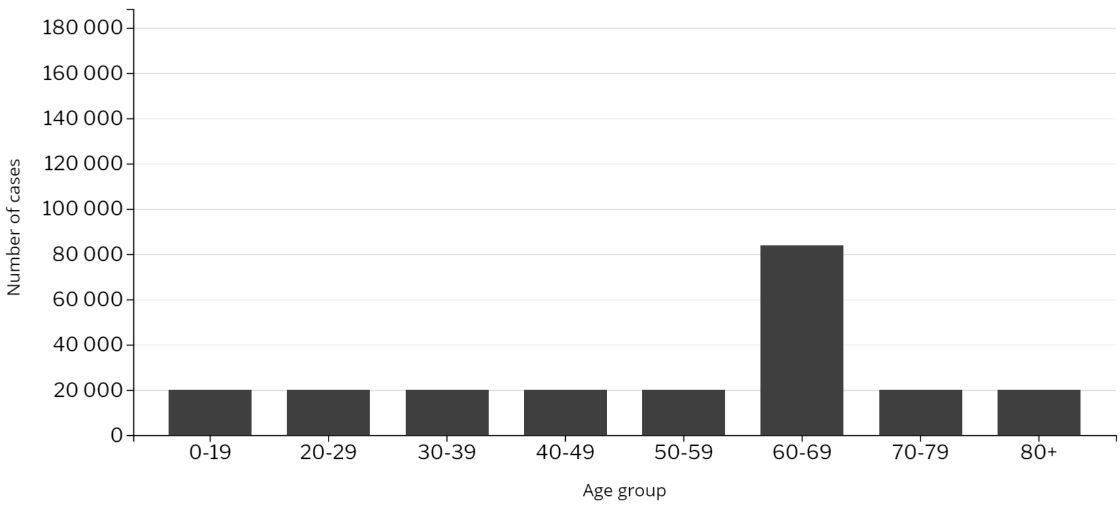

Figure 5.1: As a hint, the user is provided with the correct height of one bar.

5.4.2 Colors

The use of red and green for the colors of the arrows was meant to indicate an underestimation and overestimation respectively. After the first individual runthrough, none of the participants mentioned that the colors were strange or misleading. Even after specifically asking their opinion on it, they all agreed that the colors were a fitting choice.

To make sure that all participants really understood the meaning of the colors, they were asked to indicate on the visualisation what their best guess was. The correct answer to this question was to pick the bar with the shortest arrow independent of color. Two out of the five participants (F1, M2) immediately went for the bar with the shortest green arrow even though there was an even shorter red arrow. When asked to explain what the colors meant in this scenario, they correctly stated over and underestimation and corrected their answer. This indicates that even though people might understand the meaning of the colors when specifically asked, some subconsciously still associate green with positive and red with negative, making them draw incorrect conclusions.

When asked about possible different options for the colors, two of the five participants (M1, M3) would still keep red and green while the three others would prefer one color for all arrows over any combination of two different colors. Despite these results, we chose for two other neutral colors: blue and orange (Figure 5.2). We made this choice based on the work of Mackinlay [19] that showed that color is a very effective encoding method for categorical data, which in this case is whether the user's estimate was an over or underestimation. This information is an important part in showing the gap between the estimate and the correct data.

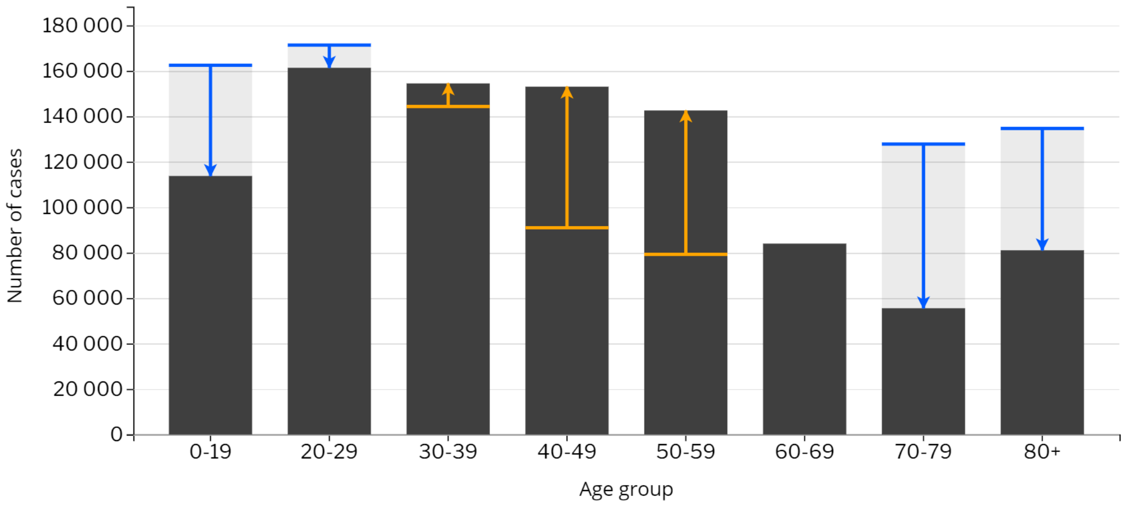

Figure 5.2: The adjusted comparison chart with blue and orange arrows.

We chose blue and orange because firstly, they don't have a negative and positive connotation like red and green but they can still be seen as a scale for higher and lower in, for example, a temperature scale (orange representing warmer and higher versus blue representing colder and lower). Secondly, these colors are easily distinguishable for almost all people who suffer from color blindness. Figure 5.3 shows how people with deuteranopia (green-blind) see the visualisation. This is the most common form of color blindness and a accounts for 75% of all cases. [20]

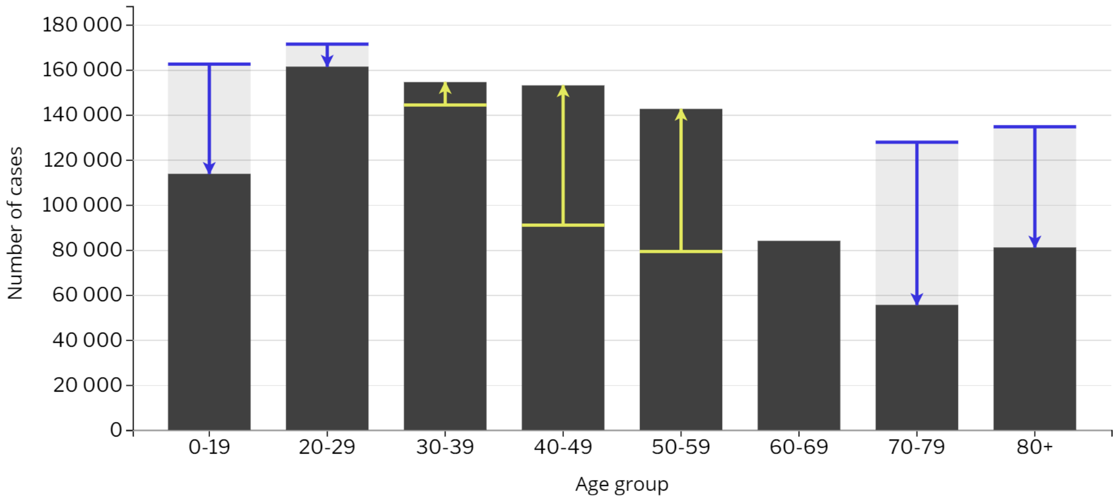

Figure 5.3: This image shows how people who suffer from deuteranopia see the visualisation.

5.4.3 Dragging and Hovering

An element that isn't in Table B.1, but was mentioned by three of the five participants (F1, F2, M2), is the fact that it wasn't immediately clear that they needed to drag the bars to estimate. For one participant (F2) it was even unclear how to drag the bars after realising that they needed to drag it. Similarly, that same participant didn't realize they could hover over a bar at the end of the animation to review the correct value and get some textual feedback. Although when pointed out, they mentioned that it didn't feel like they were missing something.

In our prototype we already changed the cursor when hovering over the top of the bar to make it clear it is draggable. Also, in the text above the chart it is mentioned that the user needs to drag the bars to change their height. Similarly, when the animation ends, the following sentence appears underneath the chart: *"Hover over the bars to study the results"*.

In the final visualisation this mention of dragging the bars was changed to a very clear instruction that stands out from the rest of the text on how to interact with the visualisation. It reads: *"Predict the values of the bars by dragging them to the desired height (do this by clicking and holding the top of each bar and dragging it)"*. The hover instruction stayed the same, but was made more visible, so it's less likely to miss.

5.4.4 Unchanged Design Elements

The *starting height* of 10 000, chosen as 10% of the maximal y-axis tick, was not viewed negatively by any of the participants. Nobody had strong opinions on any of the other options.

The *animation speed* was found too slow by one participant (M1) who would have liked the option to see them transition all at once. For this research we want to maintain consistency for all participants so we won't be including this option for the final user study. It is however something interesting to possibly include in a real world application. The other four participants found the speed ideal and would not change anything.

Regarding the animation *order* and *pauses* during the animation, none of the participants would make changes.

When *comparing with the baseline* visualisation (Figure 3.3) from the state-of-the-art [6], all five participants preferred our implementation when asked for a general preference. On the topic of more easily being able to compare the two data points, two participants (M1, F1) preferred having two bars next to each other like in the baseline visualisation. However, when reminded that the goal is to only remember the correct data points, they both decided that they preferred our implementation in this specific case.

The *dataset* was found to be interesting by all five participants, but two users (F2, M3) said they confused confirmed cases with deaths even though it was mentioned in the text above the visualisation. In the final user study a bigger emphasis was placed on making sure this was clear to the user.

Finally, when asked if they would like to use this visualisation again in the future, all five participants answered yes.

5.5 Adjusted Prototype

Based on the results of the user study, we made two changes to the visualisation. First, the correct height of one bar is provided as a hint while estimating instead of giving the total number of covid-19 cases. Second, the arrow colors red and green are replaced by blue and orange to represent over and underestimations. Any other minor changes are related to increasing the clarity of textual instructions or information. We propose the resulting visualisation (Figures 5.4 to 5.8) as the answer to RQ2: *"What bar chart visualisation is best for comparing the differences between user estimates and the correct data?"*. The appropriateness of this visualisation was further confirmed by the results of the second user study.

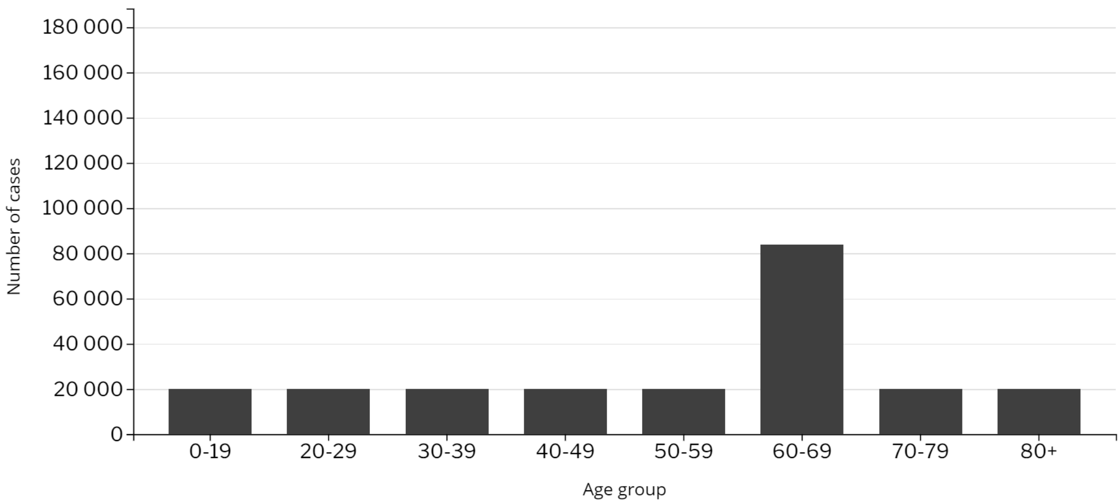

Figure 5.4: As a hint, the user is provided with the correct height of one bar.

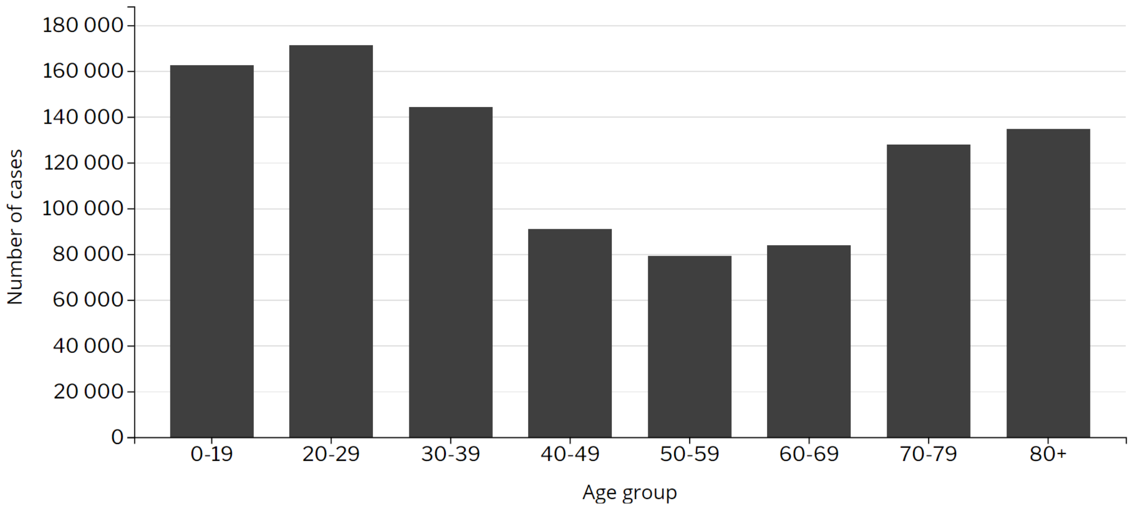

Figure 5.5: The estimation chart after a user estimated the data by dragging the bars to their desired height.

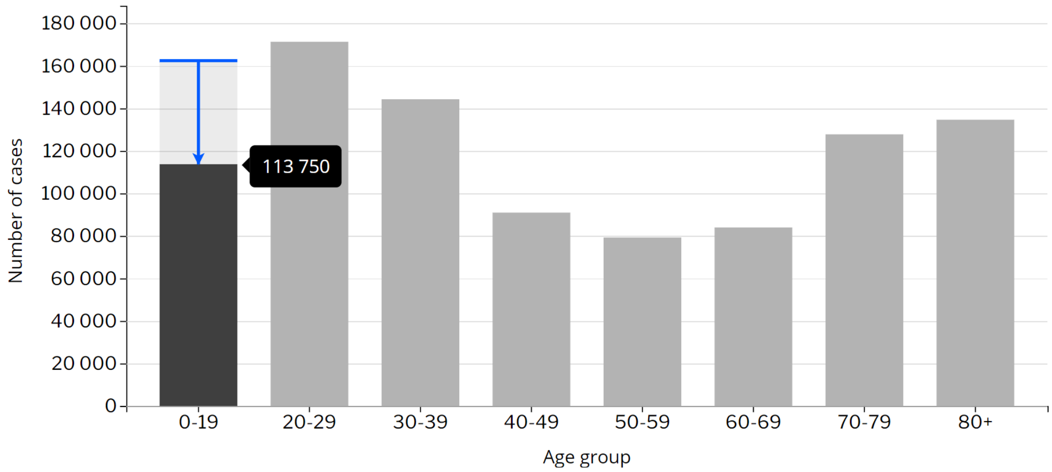

Figure 5.6: The animation starts by highlighting the leftmost bar and, in this example, shrinks the bar to the height of the correct value.

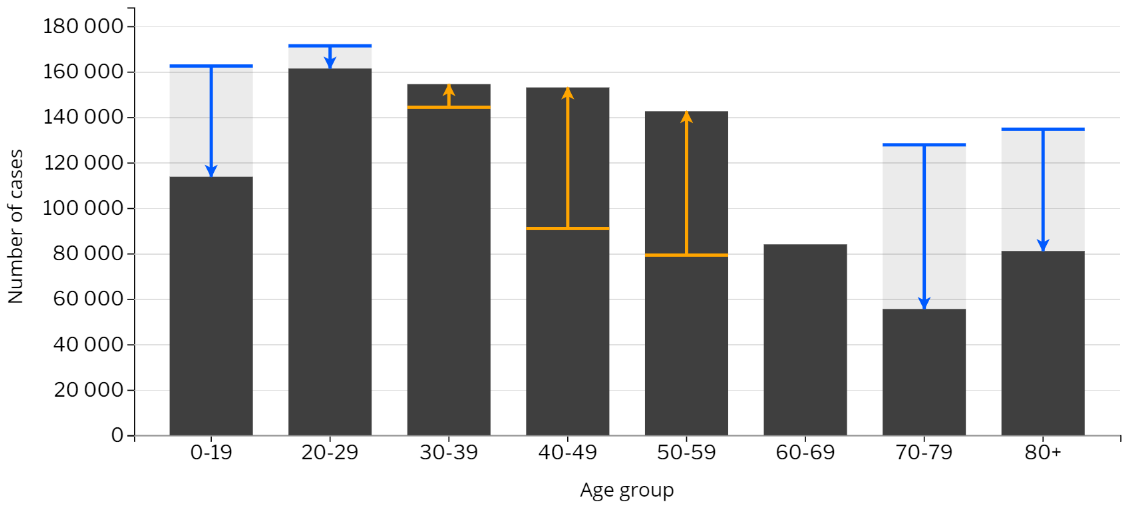

Figure 5.7: The adjusted comparison chart with blue and orange arrows.

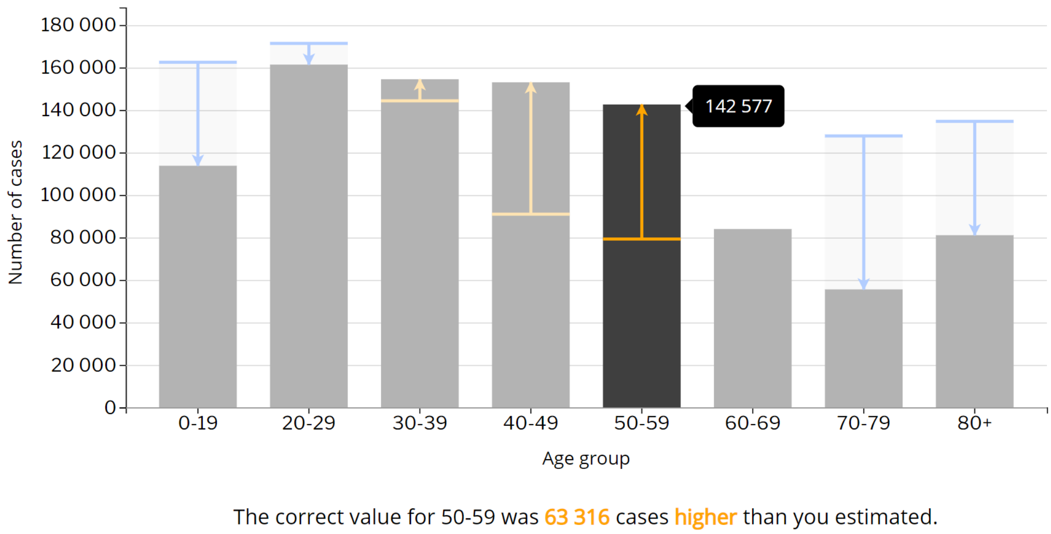

Figure 5.8: Comparison visualisation when the user hovers over a bar.

Chapter 6

Second User Study

In the second and final user study of this thesis, we determine whether eliciting and reflecting on one's prior knowledge improves data recall in bar charts. We also look at the effect of using storytelling techniques when comparing estimates with correct data on that data recall.

In Section 6.1, we relate the goals of this user study to the research questions posed in Chapter 1. The technical setup and methodology are described in respectively Section 6.2 and Section 6.3. Finally, in Section 6.4 we analyse the data that was collected during the study and report the results.

6.1 Goal

The results of the first user study produced the best bar chart visualisation to first estimate the data and to then compare those estimates with the correct data (Figures 5.4 to 5.8) as an answer to RQ2. The second user study makes use of that visualisation to find an answer to RQ1: *"What is the effect of visually eliciting and reflecting on one's prior knowledge on data recall in bar charts?"*. The evaluation of the static comparison visualisation, without added storytelling techniques, attempts to answer RQ1.1: *"What is the effect of visualising the gap between one's prior knowledge and the correct values on data recall?"*. Similarly, the evaluation of the comparison visualisation with storytelling techniques, tries to provide an answer to RQ1.2: *"What is the effect of visualising the gap between one's prior knowledge and the correct values using storytelling techniques on data recall?"*.

6.2 User Study Setup

All visualisations and pages that include those visualisations were created with Javascript with the help of the D3.js library[1], HTML and CSS. These pages were hosted on Github Pages[2]. For all pages without visualisations (introduction, asking demographic questions, etc.) we utilized the survey software platform Qualtrics[3]. Qualtrics was also used for secure data storage and evenly distributing participants over the test groups.

[1]https://d3js.org/
[2]https://pages.github.com/
[3]https://www.qualtrics.com/uk/

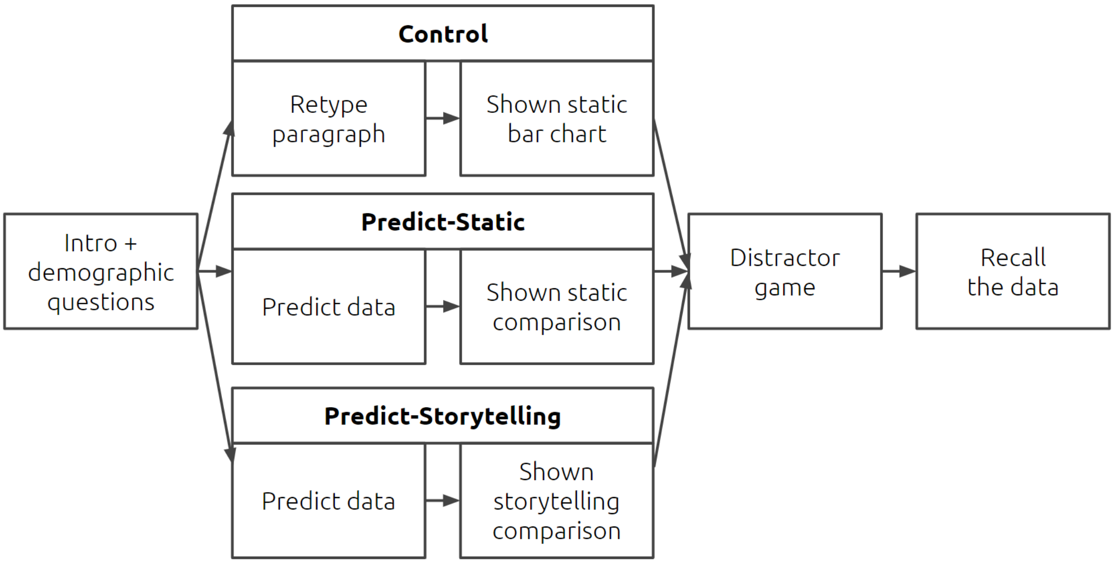

Figure 6.1: Overview of the second user study.

6.3 Methodology

Figure 6.1 shows an overview of the study methodology. This methodology is based on the between-subjects user study designed by Kim et al. [2] that is also used by Van Rooy [13]. Each step is explained in detail in the following subsections.

6.3.1 Introduction

Every participant started the user study with a short introduction. Even though the true goal of the study is to measure how well participants can recall data shown in a visualisation, to the user the goal of the study was described as: *"evaluating new and different ways to visualize data"*. This was intentionally kept vague, because if participants were informed of the true goal, they might have interacted differently with the visualisation or even tried to learn the data by heart. This would have negatively impacted the effects of any of the elicitation techniques we wanted to evaluate and render the study obsolete.

Next, participants were informed that the study would take no longer than 15 minutes and that all the data collection and processing was completely anonymous. Also, because our visualisation was not adapted for mobile devices, users were informed they could only participate on a computer or laptop. Finally, they were asked to answer four demographic questions: gender, age, highest earned degree and the amount of contact they have with data visualisations, graphs or diagrams.

6.3.2 Visualisation Interaction

From this point in the user study, participants were divided into three separate groups: a control group, a group to examine RQ1.1 called the *'predict-static'* group and another to find an answer to RQ1.2 called the *'predict-storytelling'* group. The control group only needed to observe the data and not interact with it beforehand. To ensure a similar degree of interaction with the data across all three groups, they were asked to read and retype a small paragraph about the coronavirus from the website of the World Health Organisation[4].

Both other groups needed to predict the data to elicit their prior knowledge. It was explained to them that the data was about the distribution of confirmed coronavirus cases in Belgium by age up until May 1st 2021 and that they needed to predict it. They were shown the visualisation as in Figure 5.4 and were explained that predicting could be done by dragging the top of the bars to the desired height. They were also told not to use any external tools or information to make these predictions. Predictions were adjustable and the participants could take as long as they wanted. Before they could continue, they needed to drag every bar at least once to make sure we had a prediction for each bar.

6.3.3 Visualisation Observation

The next step was to observe the data. For the control group this simply meant that they were given the correct data in a regular, static bar chart as shown in Figure 4.2. Although we label this bar chart as static, it was possible to hover over the bars to get the correct value for each bar. This was explained clearly to the user. We made this decision because this is expected behaviour from digital bar charts. To ensure consistency across all three groups, this behaviour was present in all visualisations. The users were asked to carefully examine all data at least twice before continuing.

The two predict groups needed to compare their estimates from the previous step with the correct data. For the predict-static group, the bar charts with their estimates (Figure 5.5) were instantly transformed into the static version of the comparison bar chart visualisation that was the result of the first user study (Figure 5.7). In this visualisation the bars were at their correct height and a line was drawn at the height of their estimate. There was also a colored arrow from the estimate line to the top of the bar to indicate an over or underestimation. Like all groups, they were told they could hover over the bars to see the correct values and were asked to carefully examine all data at least twice before continuing.

Where the visualisation of the predict-static group transformed instantly from the estimate to the comparison chart, the one from the predict-storytelling group transformed in an animated, bar-per-bar fashion. From left to right, one-by-one the bars grew or shrunk to their correct height, leaving a line and an arrow behind from their estimated value. When the animation finished, they ended up with exactly the same visualisation as the predict-static group (Figure 5.7) and were also told they could hover over the bars to see the correct values and were asked to carefully examine all data at least twice before continuing.

6.3.4 Distractor Task

If the participants were to have been asked to recall the data right after they just examined it, the results of the data recall task might have been skewed because they would still have had a clear image of the visualisation in their short term memory. Since we wanted to simulate the results for data recall in the long term, but actually waiting long periods of time during the user study would have been impractical, we let the users perform a distractor task.

[4]https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public

The distractor task used in this study was a 'digital paper folding test'. It was the same one that both Kim et al. [2] and Van Rooy [13] used in their papers. An example is provided in Figure 6.2. The goal of the task was to mentally visualise the folding and unfolding of a piece of paper. Every question consisted of multiple different images chronologically ordered from left to right. These images represented how the piece of paper was folded. The last image of every question showed a small circle where a hole was made through the whole folded piece of paper. The possible answers were all different images of the unfolded piece of paper with holes in different places. It was up to the user to find the correct image of the unfolded paper. The correct answer for the example was image C (Figure 6.3). There were 10 questions in total and participants had to correctly fill in as many as possible within 3 minutes.

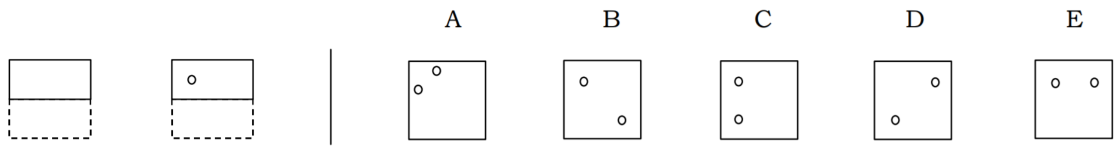

Figure 6.2: Example question of the digital paper folding test.

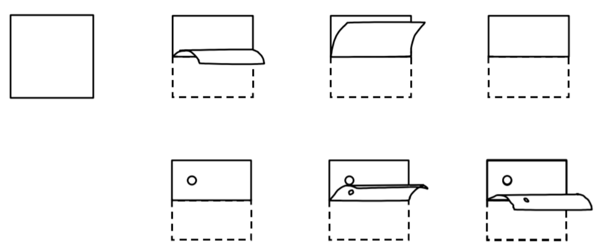

Figure 6.3: The step by step solution of the example question of the digital paper folding test.

The goal of performing this task is threefold. First, it increases the amount of time between observing the data and recalling it. Second, it prevents or weakens the mental rehearsal of the data and third, it distracts the user by giving them a new task to perform. Research has shown that these three factors actively increase the difficulty to recall previously seen information [21, 22, 23, 24].

6.3.5 Data Recall

The final step was to recall the observed data and was identical for every group. The recalling happened in the same manner as when initially estimating the data in the predict groups. The user was presented with the chart in Figure 5.4 and was asked to recall the height of all the bars except for the one that was given as a hint. They could do this by dragging the top of each bar to the desired height. Just as with the predictions, all answers were adjustable and the participants could take as long as they wanted. Before they could continue, they needed to drag every bar at least once to make sure we had a recall value for every bar.

After submitting their answers, they were shown the animated visualisation from the predict-storytelling group to show them how they performed. Afterwards, they were given the option to leave some general feedback about the study. This was optional and was mostly meant to signal the researchers if something went wrong during the survey. Finally, on the last page they were provided an explanation of the real goal of the study. They were also given the option to share the survey with other people, but were clearly asked not to mention the real goal of the study when sharing it, as to not influence the behaviour of those other participants.

6.4 Analysis and Results

6.4.1 Setup and Participant Recruitment

All data was processed and analysed using the Python programming language[5]. The Pandas library[6] was used for data preparation and processing. To make use of statistical functions we used the SciPy.stats library[7] and all charts related to the analysis were created with the Matplotlib[8] and Seaborn[9] libraries.

Before we started recruiting participants for the user study, we performed a sample size analysis based on the results found by Kim et al. [2]. The effect size between their control group and best performing test group was 0.76. A power analysis indicated that a sample size of 28 people would be needed to detect effects with 80% power using a t-test between means with alpha at .05.

The survey was distributed by privately contacting friends and relatives. Everyone over the age of 18 was allowed to participate. The study ran for 26 days and 102 users completed the survey.

6.4.2 Data Preliminaries

Of the 102 participants who completed the survey, we excluded two from our analyses. The first one had an absolute recall error of zero, meaning they recalled every of the 7 values exactly right. We assumed this was not possible without cheating so we excluded the participant in question. For the second one we had multiple indications that they did not take the survey seriously. First, they took almost 15 minutes to estimate the data, then only 18 seconds to look at the results and then took 40 minutes to recall. After all this, they were still the only participant to score worse when recalling than when estimating. After removing both participants, we ended up with 33 participants in the control group, 32 in the predict-static group and 35 in the predict-storytelling group.

The average total completion time was 11.8 minutes (SD = 4.1). There were 54 men and 46 women, the mean age was 36.1 (SD = 16.4) and the distribution of education level and amount of interaction with visualisations over all three groups can be found in Figure 6.4. The distribution of the answers to these four demographic questions for each of the three test groups separately is shown in Figure 6.5. There were no significant differences between participants in each of the groups.

[5] https://www.python.org/
[6] https://pandas.pydata.org/
[7] https://docs.scipy.org/doc/scipy/reference/stats.html
[8] https://matplotlib.org/
[9] https://seaborn.pydata.org/

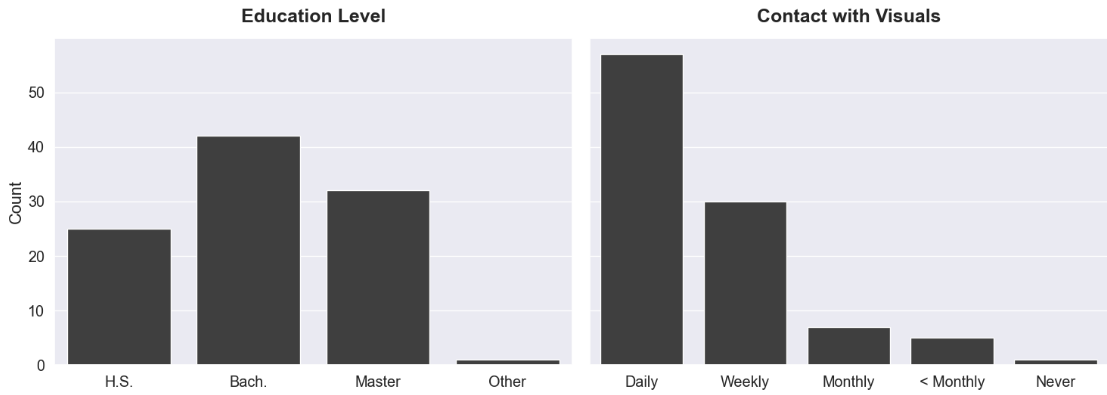

Figure 6.4: The distribution of education level and amount of interaction with visualisations over all three groups combined.

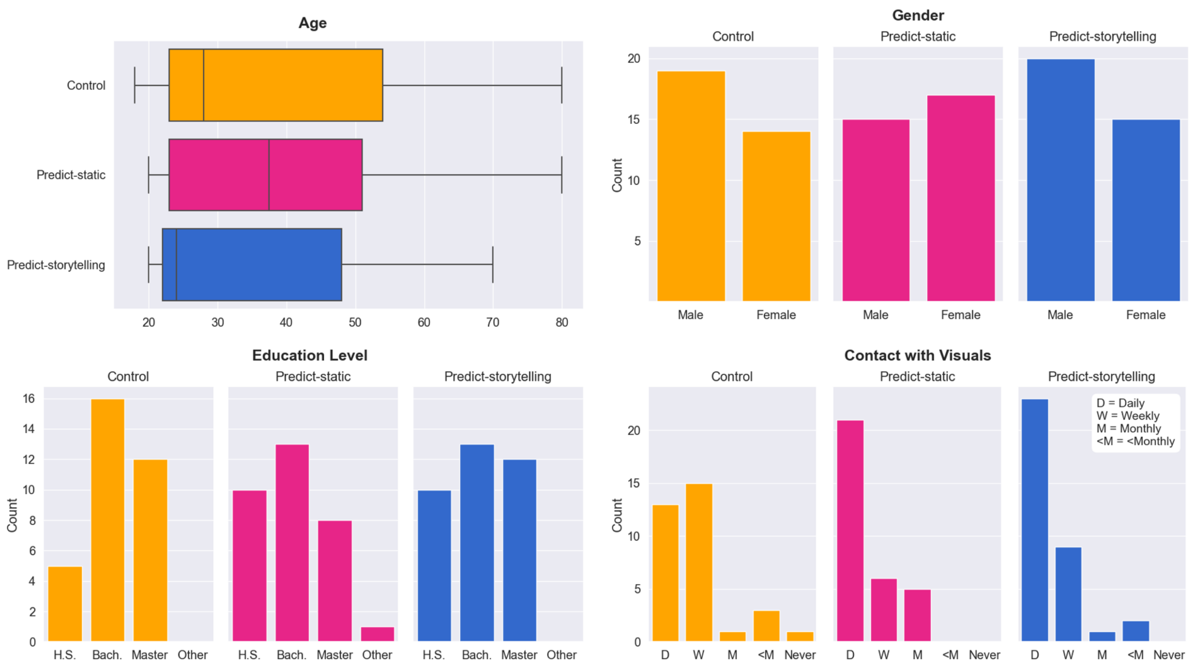

Figure 6.5: The distribution of the answers to these four demographic questions for each of the three test groups.

6.4.3 Recall Accuracy Metric

To determine the effects on data recall, we used the accuracy in recalling individual data points measured by the Average Absolute Error (AAE). All seven bars had a value ranging from 55 000 to 161 000 and all users recalled a value for each of these bars between 0 and 190 000. The absolute error for every bar for every participant was calculated by subtracting the correct value from the recall value and taking the absolute value of the result. For every participant, their AAE was calculated as the average of those seven absolute errors. A low AAE meant that the user recalled the data more accurately than participants with a higher AAE so an increase in recall accuracy is defined as a decrease in AAE, and vice versa.

Group	n	Mean	SD	Min	25%	50%	75%	Max
Control	33	21 384	11 388	1 010	14 417	20 792	28 235	52 460
Predict-Static	32	15 049	9 141	2 049	8 198	12 896	19 696	39 119
Predict-Storytelling	35	16 083	7 979	1 786	10 504	14 103	22 026	31 960

Table 6.1: Overview of the statistical values for each of the test groups

6.4.4 Results

The 32 participants who used the predict-static visualisation (μ = 15 049, SD = 9 141) recalled individual data points significantly more accurate (t(63) = 2.47, p = .016) compared to the 33 participants in the control group (μ = 21 384, SD = 11 388). Similarly, the 35 participants who used the predict-storytelling visualisation (μ = 16 082, SD = 7 979) also significantly increased data recall accuracy compared to the control group (t(66) = 2.23, p = .005). These results provide an answer to our first research question that is in line with our initial expectations: visually eliciting and reflecting on one's prior knowledge has a positive effect on data recall in bar charts.

There was, however, no significant difference in data recall accuracy between participants who used the predict-static visualisation and participants who used the predict-storytelling visualisation (t(65) = -0.49, p = .062). This result indicates that drawing the user's attention to the gap between his prior knowledge and the actual data with storytelling techniques, does not improve data recall compared to the use of a static comparison visualisation. This is not in line with our expectations for RQ1.2.

The AAE values for each of the three groups were normally distributed, as assessed by Shapiro-Wilk's test for the control group (W = .97, p = .57), predict-static group (W = .95, p = .1), and predict-storytelling group (W = .96, p = .16). There was also homogeneity of variances as determined by a Fligner-Killeen test for all participants (χ^2 = 1.37, p = .5).

A violin plot of the results can be found in figure 6.6 and an overview of the statistical values for each of the groups is shown in Table 6.1. Based on the means of all groups, there was a 29.6% increase in recall accuracy for the predict-static group and an increase of 24.8% for the predict-storytelling group compared to the control group. We believe that these increases are significant enough to have a practical use in real-life applications.

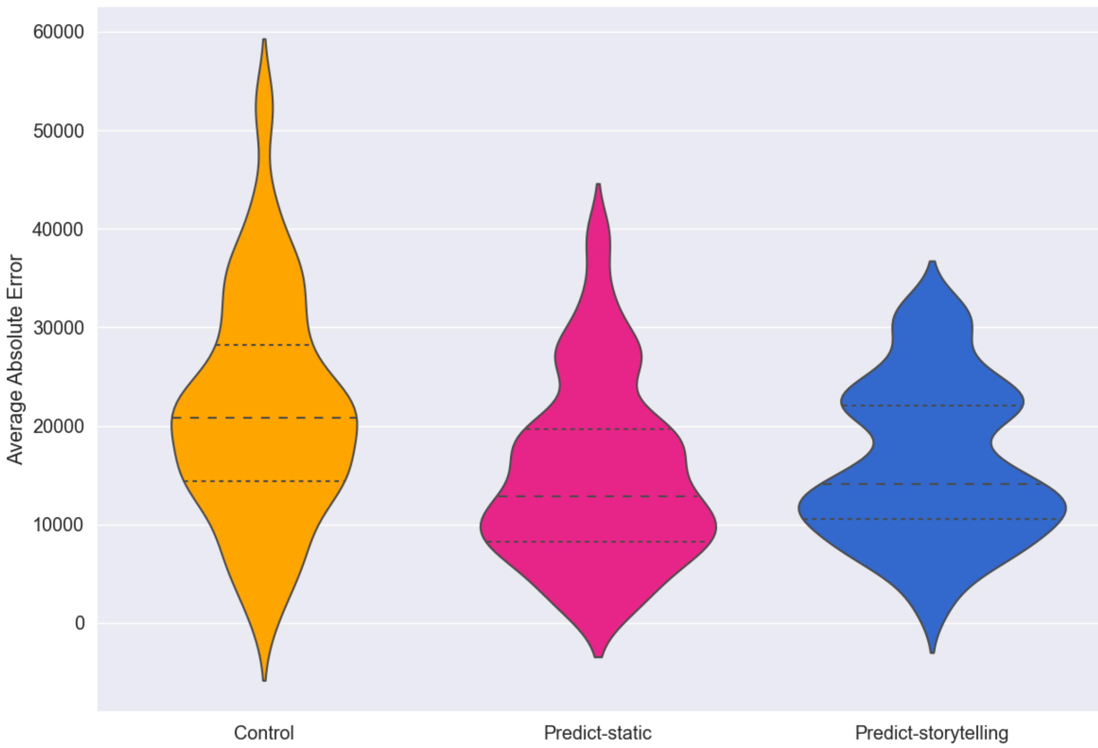

Figure 6.6: Violin plot of the AAE values for each of the test groups.

Chapter 7

Discussion

The two research questions proposed in Chapter 1 were answered by conducting two user studies. In Section 7.1 we discuss and interpret the results of both those studies and how they relate to the research questions. Section 7.2 talks about some possible practical applications and the creation of an online tool to generate belief-driven bar charts. In Section 7.3 we discuss the limitations of the study and some options that could be explored in future works.

7.1 Interpreting the Results

7.1.1 First Research Question

Kim et al. [2] found significant evidence that eliciting and reflecting on one's prior knowledge improves data recall. They however only tested this in slope charts, a visualisation specifically designed to support data with a particular structure that was beneficial for the study. When describing future work, they asked the question whether their results could be generalised to different visualisations and estimation methods. Van Rooy [13] tested this, by replicating the study for discrete geospatial visualisations. He found no significant evidence that eliciting and reflecting on one's prior knowledge in discrete geospatial visualisations improved data recall. The question now became whether the results of Kim et al. [2] were not generalisable to different visualisations or that they were generalisable but discrete geospatial visualisations were an exception to this rule.

The results of our second user study determine it is the latter. They provide significant evidence that visually eliciting and reflecting on one's prior knowledge has a positive effect on data recall in bar charts. Indicating that the results of Kim et al. [2] don't only apply to slope charts but are in fact generalisable to more general visualisations like bar charts.

In a replication study with different datasets performed by Kim et al. [2], they concluded that:

> *"Prediction, when combined with additional mechanisms like explicit self-explanation or feedback, can lead to lower absolute error even for very unfamiliar data where prediction might be difficult. [...] we saw no replication of effects for any elicitation technique on the high familiarity dataset."* [2]

We did not perform replication studies in this thesis, but based on the results of Kim et al. [2], we expect that the results of our study are generalisable for any non-high familiarity dataset.

7.1.2 Addition of Storytelling Techniques

The use of animations to draw the user's attention to the gap between their estimate and the correct data was something that Kim et al. [2] didn't test in their research but proposed when discussing future work. We expanded this use of animations to the addition of storytelling techniques and proposed RQ1.2 to test whether this would lead to a further improvement in data recall.

Our results show that the data recall accuracy of the predict-storytelling group improved significantly over the control group, but that this improvement was lower than that of the predict-static group compared to the control group. Although this difference wasn't significant, it contradicts our expectation.

Even though we did not find an increase in data recall accuracy, adding storytelling techniques might add additional benefits that were not researched in this thesis. Based on prior research [16, 17], we hypothesise that adding storytelling techniques might increase user engagement with the visualisation. In practical applications this can be important, since users who are less engaged will likely examine the data less carefully which might have a negative effect on recalling and understanding the data.

A second hypothesis theorizes that using storytelling techniques reduces the cognitive load on the user. The animations guide the user's attention through the visualisation, so they don't have to find all relevant information manually. In practical applications this potential reduction in cognitive load might prevent the user from clicking away due to mental fatigue when being presented with the results.

7.1.3 Second Research Question

The elicitation techniques and comparison methods used for slope charts and choropleths created by Kim et al. [2] and Van Rooy [13] respectively, were not generalisable to bar charts. The creation of a bar chart visualisation that could elicit prior knowledge and compare that prior knowledge with the correct data was therefore not trivial. This is why we proposed RQ2: *"What bar chart visualisation is best for first estimating data and then comparing the estimates with the correct data?"*.

Data comparison is a well researched topic in information visualisation, with multiple papers focussing specifically on bar charts. However, all these papers focussed on comparing data series of equal importance. In this research users needed to compare their estimate of the data to the correct values. These are not of equal importance, since the user should only remember the latter.

We started by determining the best comparative layout for bar charts for data series of equal importance according to the state-of-the-art [6] (Figure 3.3). We then devised some design decisions to adapt the visualisation to elicit user estimates and compare those with the correct data. These design decisions were evaluated with a think aloud study with five participants. The results of that study were very positive. Users liked the idea of first estimating the data, found it intuitive to interact with the visualisation and had no problems interpreting the comparison visualisation. When asked whether they would like to use this interactive visualisation more in the future as a possible replacement for static bar charts in online articles, they unanimously agreed.

Based on the feedback of the participants, we made two changes to the prototype and the resulting visualisation was proposed as an answer to RQ2. The effectiveness of our design was confirmed by the results of the second user study, where it was used to prove that eliciting and reflecting on one's prior knowledge improves data recall in bar charts compared to a control group.

7.2 Practical Applications of the Results

7.2.1 Practical Applications

There are multiple different reasons why authors might include data visualisations in (online) articles. When talking about climate change, they might include a chart depicting the sea levels of the last ten years to convince the reader that the ice caps are melting. Or it might simply be to give context when, for example, adding a chart that compares the sugar contents in soda's to the recommended daily sugar intake when discussing obesity in Belgium.

What both of these cases have in common, is that the author feels the data in the visualisation adds value to the article and is important enough to deserve the reader's attention. It is therefore vital to the author, as well as the reader, that the reader examines and understands the data in these charts to fully understand the point that the author tries to make.

The results in our study show that eliciting and reflecting on prior knowledge using our bar chart visualisation increases data recall, and by extension data comprehension, up to 29%. In addition, it might increase user engagement and reduce the cognitive load on the reader. We therefore propose to replace any static bar chart in online articles, with a version of our visualisation that lets the user reflect on their prior knowledge and increases their recall and understanding of the data.

Education is another field that could benefit greatly from our visualisation, when in some instances, recalling the correct data is even a goal in and of itself. With education moving further into the digital space, there is no better time than now to update the static visualisations that have been used for decades to interactive and animated visualisations that improve the learning experience and the ability to absorb the subject matter.

7.2.2 Online Tool Development

Before it's possible to introduce these kinds of visualisations into the mainstream, there needs to be an accessible way to create them. It should be simple enough so that anyone can create them while not taking longer than creating a static bar chart would. With these requirements in mind, we created an online tool with which anybody can create a belief-driven bar chart with their own data and embed it into any website or share it over the internet.

The beta version of the tool is accessible on Github Pages[1]. On the website, there are fields for the user to input the data they want in the chart. This includes the axis labels, what bar you want to give as a hint and the actual data in the form of category names with their respective values. While inputting the data, the user can preview and try out the visualisation for themselves. Once finished, they simply copy the line of code at the bottom of the page and paste it into their website's source code where they want the visualisation. It is also possible to get a link to the created visualisation that the user can share online.

7.3 Limitations and Future Work

It is now proven that eliciting prior knowledge and reflecting on it improves data recall in slope charts [2] and bar charts, but not in discrete geospatial visualisations [13]. Future work should build upon the design space for visual prediction and reflection techniques in data visualisations and evaluate whether our results are generalisable for other types of visualisations.

The dataset that was used in both user studies was deliberately chosen to make the recall task sufficiently challenging. Future work should examine the effects on the results when adjusting parameters like the familiarity of the dataset or the number of data points to make the recall task more or less challenging.

Finally, we found no increase in data recall accuracy when adding storytelling techniques to the comparison between user estimate and correct value compared to a static comparison visualisation. It is possible that the storytelling techniques we added were not sufficient in drawing the users' attention to the gap. Therefore, more research should be done on how to draw a user's attention to specific elements or pieces of information in a visualisation with storytelling techniques. We, on the other hand, hypothesize that storytelling techniques might add additional benefits to these types of comparison visualisations. Potential benefits like increased user engagement and a decrease in cognitive load could indirectly have a positive effect on data recall and data comprehension when used in practical applications like online articles or education. Studying these potential benefits and their effect on data recall is an interesting line of research for future work.

[1] jonathandelm.github.io/ShowMeYourKnowledge

Chapter 8

Conclusion

Data visualisations were designed to help users understand data better. The goal of this thesis was to increase users' recall, and by extension understanding, of data in bar charts by letting them first estimate the data and then compare their estimates with the correct data. Based on this goal, the first research question was formulated:

> RQ1: What is the effect of visually eliciting and reflecting on one's prior knowledge on data recall in bar charts?

In addition, we wanted to further increase data recall by adding additional focus to the gap between user estimate and the correct data when the user compares the two. This was done once in a static manner and once by adding storytelling techniques, resulting in two sub-questions for the first research question:

> RQ1.1: What is the effect of visualising the gap between one's prior knowledge and the correct values on data recall?
>
> RQ1.2: What is the effect of visualising the gap between one's prior knowledge and the correct values using storytelling techniques on data recall?

Before being able to evaluate RQ1 and its two sub-questions, we required a bar chart visualisation that could elicit user estimations and compare them with the correct data. We proposed our second and final research question to find this visualisation:

> RQ2: What bar chart visualisation is best for comparing the differences between user estimates and the correct data?

This visualisation was created by making adjustments to the best comparative layout for bar charts according to the state-of-the-art [6] and adding storytelling techniques to answer RQ1.2. The visualisation was evaluated in a user study that consisted of semi-structured interviews with five participants. Based on the results of the user study, two final changes were made that resulted in the visualisation that was the answer to RQ2.

With this visualisation, we could now conduct the second user study to answer RQ1. It was a between-subjects study where the 102 participants were divided into one of three groups: a control group who only saw the data in a regular bar chart or one of the groups that first had to predict the data and then either compare it in a static visualisation or one with storytelling techniques.

The study showed that the use of our belief-driven bar chart visualisation, increases data recall compared to the use of a regular, static bar chart. This answers RQ1. The results also indicated that the use of storytelling techniques when comparing between user estimates and the correct data does not improve data recall compared to a version without storytelling techniques. Meaning that there is no significant difference in the answers to RQ1.1 and RQ1.2.

Although the use of storytelling techniques does not result in an increase in data recall, we expect it might increase user engagement and reduce the cognitive load on the user. These additional benefits might indirectly result in an improvement in data recall in practical applications. Future work should be done to confirm these hypotheses.

To conclude, this research has shown that visually eliciting and reflecting on one's prior knowledge in bar charts, increases recall, and by extension comprehension, of data. The belief-driven bar chart visualisation that was developed in this thesis can be considered as a replacement for static bar charts in education and digital media.

Appendices

Appendix A

Dataset considerations and requirements

Datasets	Number of checked requirements	Relevant for target audience	Not too difficult	Not too obvious	Known on the same level by everyone	7-9 bars	Ordinal data	No need to define brackets	No focus on one bar	Best represented by bar chart	No line break needed
Confirmed coronavirus cases in Belgium by age	10	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Sugar consumption in UK	9	✓	✓	✓		✓	✓	✓	✓	✓	✓
kg GHG-emissions per kg of product with similar protein content	8	✓	✓	✓	✓	✓		✓		✓	✓
Smartphone ownership with teens	8	✓	✓			✓	✓	✓	✓	✓	✓
EU+ Total Asylum Applications	8	✓	✓	✓	✓	✓	✓	✓			✓
Percentage of content that are ads	8	✓	✓	✓	✓	✓		✓		✓	✓
Wealth disparity with the ultra rich	8	✓	✓	✓	✓	✓	✓		✓		✓
Market share of mobile gaming	7	✓	✓			✓	✓	✓	✓		✓
Chance of dying of corona by age group	7	✓	✓		✓	✓	✓			✓	✓
Annual average expenditure per household in Belgium	7	✓		✓			✓	✓	✓	✓	✓
Population of Belgium by age group.	7	✓		✓		✓	✓		✓	✓	✓

Table A.1: All considered datasets with what requirements they fulfilled.

Appendix B

Design Space for Prototype

Design Areas	Design Elements		
Estimation	Hints	*Questions*	What did you think of the hint? Was it useful?
		Oral options	Total number of cases (current)
			No hints
			Average number of cases
			Give a number of bars at the beginning
			Give the lowest and/or highest bar
	Starting height	*Questions*	What did you think of the starting height of the bars?
			Was it obvious you needed to drag the bars?
		Oral options	10% of max tick (current)
			At 0
			At the average
Animation	Speed	*Question*	What did you think about the speed of the animation?
		Oral options	Faster
			Slower
	Ordering	*Question*	What did you think of the order in which the bars were animated?
		Oral options	Left to right (current)
			Highest error first
			Lowest error first
	Pauzes	*Question*	Did you at any point want to pauze the animation?
		Oral options	One animation (current)
			Click next after every bar

Static visualisation	Comparison	*Questions*	What bar did you guess the best?
			What bar did you guess the worst?
			How did you come to these conclusions?
		Visual option	Estimation lines + arrows (current)
			Grouped bar chart with explicit encoding of error delta
	Colors	*Questions*	What does the green arrow mean?
			What does the red arrow mean?
			Did you like these colors or would you pick other ones?
		Visual options	Red/green (current)
			Non-red/green
Interactivity + messaging	Info on the chart	*Question*	Is there any extra info you wish was shown on the chart?
			If so, where would you place that extra info?
		Oral options	The value you estimated
			By how much you were off
			By how much % you were off
	Info as text	*Questions*	Did you read the text under the chart when hovering?
			Did you think this was useful to read?
			What extra info, if any, would you like to see there?
		Oral options	The value you estimated
			By how much you were off
			By how much % you were off
			What bar was your best and/or worst
			Info about your overall performance (average, ...)

Dataset	Dataset	*Questions*	What did you think about the dataset? Was it interesting data? Before guessing, how confident were you about your guess? Did you think it would be too hard/too easy? Were you surprised by the differences between your guesses and the results? After seeing the results, did you still think it was too hard/too easy?
General	General	*Questions*	What is your overal feeling about this visualisation? Would you like to use this visualisation more in the future?

Table B.1: A list of all design areas of our prototype with their specific design elements, what questions we asked the user in the first user study and, if relevant, some different design options that were considered during development.

Appendix C

GDPR Information Sheet

KU LEUVEN

GDPR information sheet

INFORMATION ON THE PROCESSING OF YOUR PERSONAL DATA

As a result of your participation in the interactive bar charts study, personal data relating to you will be collected and processed. These data will be processed in accordance with the General Data Protection Regulation (GDPR). With this information sheet, we would like to inform you about the use and storage of your data.

In particular, the following personal data will be collected during the study: video and audio recordings.

Use of your personal data

Only personal data required for the purposes of this study will be collected and processed. More specifically, the study aims to study interactivity, animations and storytelling techniques in bar charts. The collected data may possibly be re-used in future studies.

Data collected for this study will be pseudonymised. This means that data that might identify you, such as video and audio recordings, will be separated from the other data in the study and replaced by a unique random code. In this way, the data can no longer easily be attributed to a specific data subject. Only the researcher can link the data to a specific individual by means of the unique code. However, this will only happen in exceptional cases, for example if you wish to exercise your right to access, rectify or erase your data. You will also not be identified in publications arising from the research.

The data will be processed on the basis of public interest. This means that the research will lead to advances in knowledge and generate insights that (directly or indirectly) benefit society.

Your data will be stored by the researchers for 10 years after the end of the study at a secure storage location at KU Leuven. After this period, your personal data will be permanently deleted if they are no longer needed for the purposes of the research.

Your rights

You have the right to request more information about the use of your data. In addition, you have the right to access, rectify or erase your data unless exercising these rights would render impossible or seriously impair the achievement of the research objectives.

If you wish to exercise one of these rights, please contact the researchers using the contact details on the next page of this information sheet.

Figure C.1: GDPR Information Sheet (1/2)

KU LEUVEN

GDPR information sheet

Contact details

For the purposes of this research, KU Leuven is the data controller. More specifically, only the researchers involved Jonathan Delmeiren, Diego Rojo Garcia and Katrien Verbert will have access to your personal data. Should you have any specific questions about this study, including the processing of your personal data, please feel free to contact them.

Jonathan Delmeiren: j.delmeiren@gmail.com

For any further questions and concerns regarding the processing of your personal data, please contact Toon Boon, KU Leuven's data protection officer for research (dpo@kuleuven.be) . Please specify the study concerned by mentioning the title as well as the names of the researchers involved.

If, after contacting the data protection officer, you would still like to lodge a complaint about the use of your personal data, you can contact the Belgian Data Protection Authority (www.gegevensbeschermingsautoriteit.be).

Figure C.2: GDPR Information Sheet (2/2)

Appendix D

Introduction First User Study

Before sending Microsoft teams link

Tell them not to use their real name and just put 'User'.

Introduction

My name is Jonathan Delmeiren and for my master's thesis in Applied Informatics, I'm designing a new, interactive visualisation of bar charts to improve user engagement. The setting for this visualisation would be in online articles or papers where they use bar charts to show their data. There are some other goals that I will be researching, but for this interview it may be best not to have too much information beforehand. If you're interested, I can give you a more in depth explanation after the interview is completed.

More practically, this interview should not take longer than 60 minutes and is entirely voluntary. So if you at any point want to stop the interview, that is totally fine. Also, all your answers will be anonymous, so you won't be mentioned anywhere in the thesis itself. If it's okay for you, I will also be recording this interview to more easily analyse the results afterwards. I will be the only person that reviews these videos. I will also send you the GDPR information sheet. You don't have to sign anything, it just lists what data will be collected (video and audio), how it's stored and who to contact if you have any questions about it afterwards.

Last but not least, it is very important for you to know that there are no wrong answers in this interview. Everything you say is valid and useful, so please don't hold back, since the more feedback you give, the better the interview is for me. Okay, if you don't have any questions, I will start the recording and the interview itself.

Demographic questions

I will start off by asking you some basic demographic questions to include in the analysis afterwards.

- What is your gender?
- What is your age?

- What is your highest earned degree?
- How much do you come in contact with visualisations like bar charts? (per day/week/month)
- How much do you read online articles? (per day/week/month)

Interview

Now how is the interview going to go? I'm going to give you a link to a website that has the visualisation inside an article. For the first part I just want you to go through everything on the page like you came across this on the website of an online paper and you're just interested in reading this article. But while you are reading and interacting, I want you to say everything you think out loud. This can really be anything. Things you don't understand or didn't expect or just things you find ugly or pretty or good or bad. Literally anything, the more the better. During this time I won't be asking you any questions and interrupt as little as possible. Then afterwards, we will go over everything again but then I will ask you questions about specific things. Okay, here is the link (and remember, just say anything you think of): *link*

Bibliography

[1] J. Katz, "You draw it: Just how bad is the drug overdose epidemic," *The New York Times*, 2017.

[2] Y.-S. Kim, K. Reinecke, and J. Hullman, "Explaining the gap: Visualizing one's predictions improves recall and comprehension of data," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 1375–1386, 2017.

[3] S. LYi, J. Jo, and J. Seo, "Comparative layouts revisited: Design space, guidelines, and future directions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1525–1535, 2020.

[4] haikuanalytics.com, "Excel tricks for data viz," 2021. http://haikuanalytics.com/excel-tricks-for-data-viz-1/.

[5] G. Carenini, C. Conati, E. Hoque, B. Steichen, D. Toker, and J. Enns, "Highlighting interventions and user differences: Informing adaptive information visualization support," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1835–1844, 2014.

[6] A. Srinivasan, M. Brehmer, B. Lee, and S. M. Drucker, "What's the difference? evaluating variations of multi-series bar charts for visual comparison tasks," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[7] F. Nguyen, "Belief-driven data journalism," *Computation+ Journalism*, 2019.

[8] M. Hegarty, "Diagrams in the mind and in the world: Relations between internal and external visualizations," in *International Conference on Theory and Application of Diagrams*, pp. 1–13, Springer, 2004.

[9] J. G. Trafton, S. B. Trickett, and F. E. Mintz, "Connecting internal and external representations: Spatial transformations of scientific visualizations," *Foundations of Science*, vol. 10, no. 1, pp. 89–106, 2005.

[10] Z. Liu and J. Stasko, "Mental models, visual reasoning and interaction in information visualization: A top-down perspective," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 6, pp. 999–1008, 2010.

[11] E. Stern, C. Aprea, and H. G. Ebner, "Improving cross-content transfer in text processing by means of active graphical representation," *Learning and Instruction*, vol. 13, no. 2, pp. 191–203, 2003.

[12] R. Cox, "Representation construction, externalised cognition and individual differences," *Learning and instruction*, vol. 9, no. 4, pp. 343–363, 1999.

[13] M. Van Rooy, "Show me your knowledge: Using prediction to improve data recall on geospatial visualisations," 2020.

[14] B. Ondov, N. Jardine, N. Elmqvist, and S. Franconeri, "Face to face: Evaluating visual comparison," *IEEE transactions on visualization and computer graphics*, vol. 25, no. 1, pp. 861–871, 2018.

[15] E. Segel and J. Heer, "Narrative visualization: Telling stories with data," *IEEE transactions on visualization and computer graphics*, vol. 16, no. 6, pp. 1139–1148, 2010.

[16] Q. Zhi, A. Ottley, and R. Metoyer, "Linking and layout: Exploring the integration of text and visualization in storytelling," in *Computer Graphics Forum*, vol. 38, pp. 675–685, Wiley Online Library, 2019.

[17] S. McKenna, N. Henry Riche, B. Lee, J. Boy, and M. Meyer, "Visual narrative flow: Exploring factors shaping data visualization story reading experiences," in *Computer Graphics Forum*, vol. 36, pp. 377–387, Wiley Online Library, 2017.

[18] G. A. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information.," *Psychological review*, vol. 63, no. 2, p. 81, 1956.

[19] J. Mackinlay, "Automating the design of graphical presentations of relational information," *Acm Transactions On Graphics (Tog)*, vol. 5, no. 2, pp. 110–141, 1986.

[20] colourblindawareness.org, "Types of colour blindness," 2021. https://www.colourblindawareness.org/colour-blindness/types-of-colour-blindness/.

[21] D. Laming, "Analysis of short-term retention: Models for brown-peterson experiments.," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 18, no. 6, p. 1342, 1992.

[22] T. J. Ricker, E. Vergauwe, and N. Cowan, "Decay theory of immediate memory: From brown (1958) to today (2014)," *Quarterly Journal of Experimental Psychology*, vol. 69, no. 10, pp. 1969–1995, 2016.

[23] J. Brown, "Some tests of the decay theory of immediate memory," *Quarterly journal of experimental psychology*, vol. 10, no. 1, pp. 12–21, 1958.

[24] L. Peterson and M. J. Peterson, "Short-term retention of individual verbal items.," *Journal of experimental psychology*, vol. 58, no. 3, p. 193, 1959.

AFDELING
Straat nr bus 0000
3000 LEUVEN, BELGIË
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
www.kuleuven.be