FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Intuitive Visualization of Temporal Uncertainty

DIPLOMARBEIT

im Rahmen des Studiums

Visual Computing

eingereicht von

Fabian Schwarzinger
Matrikelnummer 1225307

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung: Univ.Prof.in Mag.a Dr.in Silvia Miksch
Mitwirkung: Univ.Ass.in Mag.a Dipl.Ing.in Dr.in Theresia Gschwandtner

Wien, 6. August 2018

(Fabian Schwarzinger) (Silvia Miksch)

FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

Intuitive Visualization of Temporal Uncertainty

DIPLOMA THESIS

in

Visual Computing

by

Fabian Schwarzinger
Registration Number 1225307

to the Faculty of Informatics
at the Vienna University of Technology

Advisor: Univ.Prof.in Mag.a Dr.in Silvia Miksch
Assistance: Univ.Ass.in Mag.a Dipl.Ing.in Dr.in Theresia Gschwandtner

Vienna, 6th August 2018

(Fabian Schwarzinger) (Silvia Miksch)

Erklärung zur Verfassung der Arbeit

Fabian Schwarzinger
Karolinengasse 6/10, 1040 Wien

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

(Ort, Datum)	(Fabian Schwarzinger)

Acknowledgments

First and foremost, I would like to thank my parents for their continuing support and for enabling my studies in the first place. I would also like to thank Silvia Miksch for advising this thesis and express special thanks to Theresia Gschwandtner for her stellar assistance of my work. This thesis would not be what it is without her encouraging words, constructive feedback and exceptional commitment. Furthermore, I would like to thank Marlies Peter for proofreading my work.

Vital for this work is also the time committed by all study participants; many thanks to each of them. Last but not least, I want to thank all my friends and especially my partner for keeping me motivated throughout my studies.

Abstract

Information retrieved from the real world often contains some kind of inherent uncertainty. In recent years there has been an effort to incorporate this aspect of data into visualizations. This is also true for the representation of temporal data. How this can be done in an intuitive way is still one of many open questions regarding temporal uncertainty visualization.

This thesis presents two subsequent user studies, which aim at providing insights into this matter. In the first study, called the Drawing Study, 32 participants are asked to draw their own visualization designs, based on predefined scenarios and tasks. These drawings are analyzed through an open coding approach. The analysis yields a list of hypotheses regarding the intuitiveness of temporal uncertainty visualization. From this list a selection of hypotheses leads to more concrete research questions, which form the basis for the second study, called the User Survey. In this online survey 60 participants compare and rate different visualization approaches in several scenarios.

This rating of intuitiveness yields valuable insights for future visualization design. The results indicate that icon representations are not considered intuitive, even though they might seem to be at first glance. It can be argued that icons just need to be designed specifically for certain tasks and scenarios to be perceived intuitive. Furthermore, the results suggest that most people prefer to have uncertainty presented to them, even if it is not relevant for the task at hand. This finding could have important implications for the design of future visualizations.

Kurzfassung

Datensätze, die aus der „echten Welt“ stammen, enthalten meist verschiedene Arten inhärenter Unsicherheiten. In den letzten Jahren gab es Bemühungen, diesen Aspekt der Daten in Visualisierungen zu berücksichtigen und sichtbar zu machen. Dasselbe gilt für die Darstellung von zeitlichen Daten. Wie dies am besten auf eine möglichst intuitive Art bewerkstelligt werden kann, ist in Bezug auf die Visualisierung von zeitlichen Unsicherheiten noch eine von vielen offenen Fragestellungen .

In dieser Arbeit werden zwei aufeinander aufbauende Nutzerstudien präsentiert, welche darauf abzielen, Erkenntnisse in diesem Bereich offenzulegen. In der ersten Studie, bezeichnet als Drawing Study, wurden 32 Teilnehmer darum gebeten, basierend auf gegebenen Szenarien und Aufgabenstellungen, eigene Visualisierungsdesigns zu entwickeln und zu zeichnen. Diese Zeichnungen wurden mit einem sogenannten Open Coding-Ansatz ausgewertet. Das Ergebnis dieser Analyse ist eine Liste von Hypothesen in Bezug auf die intuitive Darstellung von zeitlichen Unsicherheiten. In der Folge wurden aus dieser Liste Hypothesen ausgewählt, welche zu genauer definierten Forschungsfragen führten. Diese Fragen bilden die Basis für die zweite Nutzerstudie, welche als User Survey bezeichnet wird. In dieser Onlineumfrage wurden 60 Teilnehmer gebeten, jeweils zwei Visualisierungstechniken in gegebenen Szenarien zu vergleichen und zu bewerten.

Diese Bewertung der Intuitivität der Techniken führte zu Erkenntnissen, welche für die Entwicklung von zukünftigen Visualisierungen wichtig sein könnten. So weisen die Resultate unter anderem darauf hin, dass symbolische Darstellungen unerwarteterweise nicht intuitiv wahrgenommen werden. Ein Grund dafür könnte sein, dass solche Darstellungen speziell auf bestimmte Einsatzgebiete zugeschnitten werden müssen, um intuitiv zu sein. Außerdem deuten die Studienergebnisse darauf hin, dass die Visualisierung von Unsicherheiten von den meisten Menschen bevorzugt wird, auch wenn die Unsicherheiten selbst nicht von Bedeutung für die vorliegende Aufgabe sind. Diese Erkenntnis könnte wichtige Implikationen auf die Entwicklung von zukünftigen Visualisierungen haben.

Contents

CHAPTER 1

Introduction

"A technical system is, in the context of a certain task, intuitively usable while the particular user is able to interact effectively, not-consciously using previous knowledge."

— Naumann, Hurtienne, Israel, Mohs, Kindsmüller, Meyer and Hußlein,
Intuitive use of user interfaces: defining a vague concept. [44, p. 129]

1.1 Motivation

Information retrieved from the real world is almost always affected by inherent uncertainty, because measurements can generally never be perfectly exact. Furthermore, data is often directly aggregated during its time of capture. The same can be said about temporal data. Exact times of events might not be known (e.g., 'time of the big bang'), they might be given in an imprecise way (e.g., 'during the past hour') or a prediction of the future, which is inherently vague (e.g., 'it will take one or two weeks'). This kind of temporal uncertainty we are concerned with in this work stands in contrast to uncertain values of time-oriented data. I. e. our focus lies on uncertain time frames and not on uncertain values in time. Furthermore, temporal uncertainty is also different from branching time models, which depict time as a graph that shows multiple possible outcomes of future events [3]. Most existing information visualization (InfoVis) systems do not incorporate uncertainty into their visual representations. Due to a shift from deterministic approaches toward statistical models in certain areas, such as business, an increasing need for novel visualizations arises [6]. To incorporate uncertainty into visual representations is the aim of many works published in recent years. Several of them specifically focus on communicating temporal uncertainty to users [22, 34, 10, 41, 4, 24].

It is important that novel visualization designs are properly evaluated. This ensures that the target user group is effectively supported in their tasks. Furthermore, evaluation should not only be done after a new approach is fully implemented, but also beforehand. This preceding evaluation aims to evaluate the need of the target user, so the visualization can be specifically designed to support this need. Tamara Munzner's Nested Model [43] gives a thorough characterization of evaluation methods. It encompasses evaluation done before, during and after the design process of novel approaches.

Not all user studies aim to evaluate specific visualization approaches, but try to answer more general questions [61, 60, 18]. These exploratory experiments serve the purpose of gaining general insights about visualizations and their usage. Furthermore, they can discover and identify potential directions for future research. Most of the empirical work done in the domain of InfoVis aims to evaluate and compare existing and newly devised designs [36]. This often leaves the question open if the evaluated design was the user's first choice, or if there is another preferred or more intuitive solution. This is why it is important to do exploratory research to gain insights toward the intuitiveness of visualizations and guide future research.

1.2 Aim of this work

The main goal of this work is to generate insights about the intuitive visualization of temporal uncertainty. Our understanding of the term 'intuitive' in this context stems from the definition of Naumann et al. who state: *'A technical system is, in the context of a certain task, intuitively usable while the particular user is able to interact effectively, not-consciously using previous knowledge.'* [44, p. 129]

To this end our research goes through two main practical steps. In the first step we want to explore this domain to generate interesting hypotheses. In this context *interesting* means that the theory's validation will lead to valuable insights that can be incorporated into future visualization designs. This means that no existing, predefined visualization designs can be used at this stage, since we want to keep the design space that is being explored as open as possible.

The first step yields a list of hypotheses which are unvalidated. From this list a selection of the most promising theories is made. In step 2 this selection is then validated to generate concrete results toward our main goal. To render these results valuable for future research in this domain, it is important to test them in a way that keeps them generalizable and therefore applicable to as many similar scenarios as possible.

Research Questions

The following list presents the main research questions answered by this thesis. They originate from the selection process of hypotheses in the second step described above.

- **RQ1** Is it more intuitive to the average user to use Gradient Plots (see Figure 2.25) or temporal line charts to judge a specific probability value of an event to a given point in time?
- **RQ2** Is it more intuitive to the average user to visualize a comparison of two events, with uncertain temporal bounds, in a superimposed view (overlapping representation) or a juxtaposed view (side-by-side view)?
- **RQ3** Is it more intuitive to the average user to use an explicit uncertainty representation or uncertainty encoded in icons if the task at hand only calls for a rough approximation of the probability?

- **RQ4** Is it more intuitive to visualize an underlying uncertainty or to omit it if the uncertainty is not directly relevant for the task at hand?

Figure 1.1 shows an overview of the two main steps of this work and their in- and outputs. How the presented goals of the separate steps are addressed is explained in the following Methodology section.

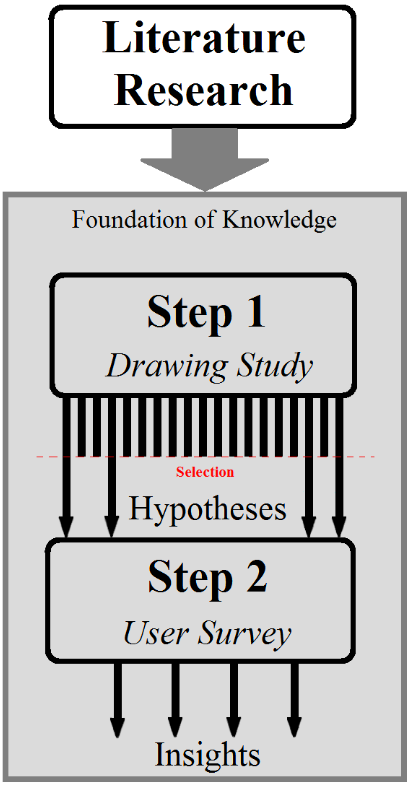

Figure 1.1: *Before any practical work is done, a comprehensive literature research is conducted to build a profound foundation of knowledge about the state of the art within the field. Building upon this knowledge the first main step is a user study called the Drawing Study (described in detail in Chapter 3). It yields a number of hypotheses, which are filtered for the most promising ones. This selection is then evaluated in a user survey to gain valuable insights about the visualization of temporal uncertainty.*

1.3 Methodology

The first part of this work consists of a comprehensive literature research. This yields the necessary foundation of knowledge to build our work upon. Firstly, it informs about the current state of research in the domain of temporal uncertainty visualization and its open questions to determine which insights would be new and valuable to the research domain. Secondly, general approaches of empirical user studies are identified to be utilized in our own study design.

Thirdly, existing user studies are analyzed in regard to how they were designed, conducted and evaluated to learn about conventions and typical approaches to empirical work within the field. The most relevant works of this literature research are presented in the State of the Art Chapter.

This first part of gathering a comprehensive knowledge foundation is followed up by the first practical step towards our goal of generating insights. It consists of an exploratory user study, which we call the Drawing Study. The goal of this step is to address the need for interesting hypotheses, as described in the previous section. Since the results of the Drawing Study leave much room for interpretation, it is possible to generate many hypotheses from them. Some of these could be of interest for future research, but not all of them are evaluated in the course of this work. Instead we select only the most promising theories for evaluation. The assessment of them happens in the second practical step, which we call the User Survey. It is an online survey, which asks participants for their opinion about presented visualizations, with the aim of answering the selected research questions. All the details about the design and conduct of this survey are presented in Chapter 4, while the results are separately presented in Chapter 5.

1.4 Structure of this thesis

This thesis is structured in the following chapters:

- **Chapter 2 State of the Art:** In this chapter we present the state of the art of temporal uncertainty visualization and other literature relevant for our work. The presented literature is split up into three main categories. The first one is about techniques for visualizing temporal uncertainty. The second one is about the theory of user study design and the third one presents existing user studies in the domain of InfoVis.

- **Chapter 3 Drawing Study:** Here we present our exploratory Drawing Study. The rationale behind its design is thoroughly explained. The results are analyzed through an open coding approach, which is also described in detail. This analysis yields a list of hypotheses about the visualization of temporal uncertainty, which are partly tackled in our User Survey.

- **Chapter 4 User Survey:** This chapter encompasses our online User Survey. The selection of hypotheses from the Drawing Study and the formulation of research questions from those hypotheses is explained. Furthermore, the study's design based on the selected goals is described in detail.

- **Chapter 5 Results:** In this chapter the results of the User Survey are presented and analyzed in regard to the posed research questions. The results are also discussed with regard to their implications and the insights they provide. Furthermore, the qualitative feedback gathered from participants, as well as the limitations of this work are discussed. The last part also encompasses possible directions for future research.

- **Chapter 6 Conclusion:** We conclude this thesis with an overview of our work and its main contributions.

CHAPTER 2

State of the Art

This chapter presents relevant literature for the design and evaluation of the practical part of this thesis. The entirety of the collected works is split up into three main groups.

Before the findings are presented, the methodology of how this literature was explored and found is explained in Section 2.1.

The first group, presented in the Section 2.2, is about the visualization of temporal uncertainty. It encompasses a multitude of techniques, starting from early approaches like the box plot [23] to more novel techniques.

In Section 2.3 the second group of literature is presented, which provides necessary theoretical knowledge about the design of user studies. There are many ways how such a study can be conducted, which becomes especially apparent if one considers small detail decisions made during the design. It might be these small, seemingly insignificant, decisions that have an important impact on the outcome of the experiment. Therefore, it is important to gather knowledge to be able to avoid pitfalls and to correctly design a successful user study.

To also see how this theory can be put into practice, a selection of existing user studies in the realm of InfoVis is presented in Section 2.4 of this chapter. Reviewing state-of-the-art experiments is an important step to learn and understand how to design and evaluate such experiments.

The following Section 2.5 is compiled as a list of the most important points of all reviewed works in regard to this thesis and the corresponding practical work. All these points are again addressed in Section 2.6. There the relevance of every point in regard to the presented practical work and how it impacts the final study design is explained in detail.

2.1 Method

To get a good overview of existing literature and to find relevant works to base this thesis on, first and foremost the *Google Scholar* [21] search engine was used. The main reason for this is that this engine allows to search through a multitude of digital libraries at once and to also conveniently find referencing relationships between works. Additionally, the *ACM Digital Library* [7], as well as *IEEE Xplore* [28] were utilized.

Since this thesis follows up on the work of Gschwandtner et al. [22], the literature research was started by reviewing publications cited by their paper and more recent work referencing it. After broadening the search to similar papers, all findings were structured into the three main categories of literature of this chapter. Subsequently, more literature was reviewed in every category.

The first category aims to give an overview of the state-of-the-art of temporal uncertainty visualization. Search queries for this category included, for example, *temporal uncertainty visualization*, *temporal confidence visualization* or *uncertainty visualization temporal data*.

The second category is about the theory of user study design in the context of visualization. To find literature in this context, we used search queries like *visualization user study*, *visualization human factors*, *visualization user centered design* and similar ones.

References for the third category which features existing user studies, from which more about the actual implementation of such can be learned, mostly emerged in the same search process as the last category. The reason for this is that most theoretical literature either reviews existing study implementations or references them as examples.

2.2 Visualization of temporal uncertainty

One of the oldest techniques that is suited for visualizing temporal uncertainty is the *box plot*. Its origins lie as far back as Haemer's work from 1948 [23], in which he describes the *range bar chart*. This representation is very similar to the first mention of the box plot by Tukey [58] and encodes the same five values that describe the uncertainty of the data of interest. As shown in Figure 2.1 it encodes the median, the upper and lower quartiles, as well as the maximum and minimum value. By aligning a box plot along a time axis, it can be used to encode the uncertain time of an event. It can also be used to represent the uncertain start and/or end times of an event.

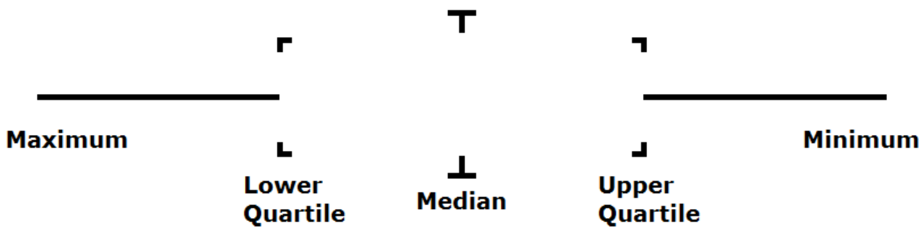

Figure 2.1: *A box plot may be presented in slightly different appearance, but it always encodes the same five values: median, upper and lower quartile and the maximum and minimum value. (original illustration)*

There are further variations of the box plot, like the *range plot* by Spear and Spear [55] or the *interquartile plot* by Tufte and Schmieg [57], which only change the visual appearance of the plot, but still encode the same five values. Examples of these can be seen in Figure 2.2.

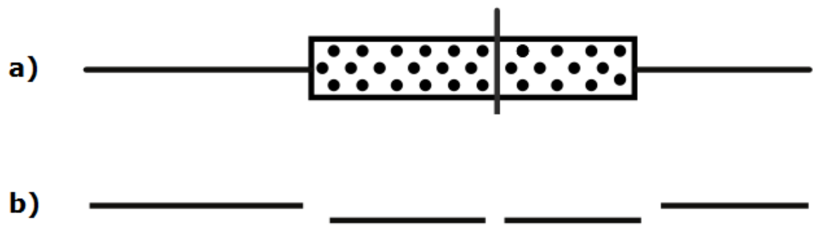

Figure 2.2: *The range plot (a) and the interquartile plot (b) are visual modifications to the box plot and also encode the same five values to describe a distribution of interest. [48]*

Besides these cosmetic modifications of box plots, there are also variations that aim to convey the underlying distribution. Typically this is done by encoding density information at the sides of the box plot. Benjamini [8] proposed the *Histplot*, which is constructed in the following way: A traditional box plot is drawn, with an additional encoding of the density at the median and both quartiles, through the width of the box. These density landmarks are connected by straight lines. Examples of the Histplot can be seen in Figure 2.3.

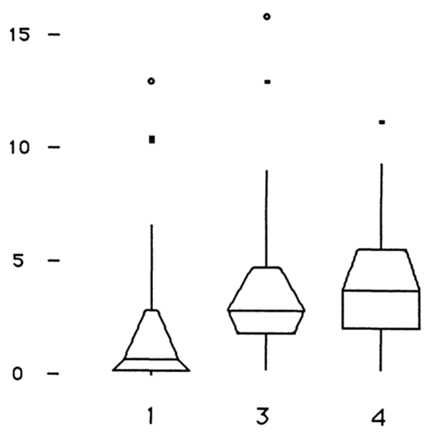

Figure 2.3: *The Histplot encodes the density of the underlying distribution at the median and both quartiles through its width. [8]*

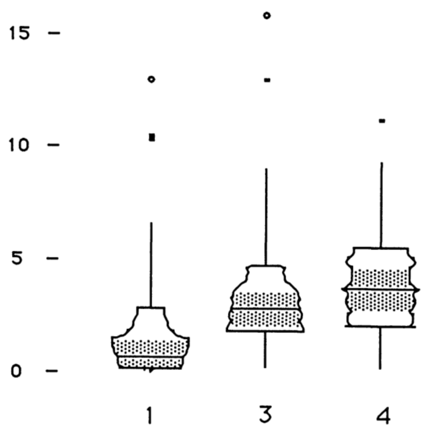

Figure 2.4: *The Vaseplot encodes the density of the underlying distribution at multiple points between both quartiles. [8]*

Furthermore, Benjamini proposed another variation, called the *Vaseplot*. It is constructed in the same way, but the density is estimated at various points within the quartiles. The result of this is shown in Figure 2.4. The final appearance of this plot is highly dependent on the amount of density estimates used to draw the plot, which is illustrated in Figure 2.5.

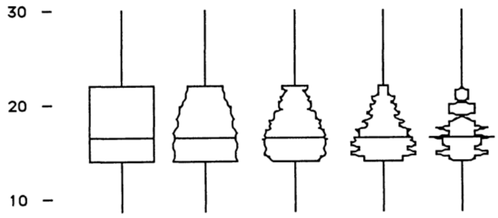

Figure 2.5: *The five presented Vaseplots represent the same underlying distribution, but with a different amount of density estimates between the quartiles, which strongly changes the appearance of the plot. [8]*

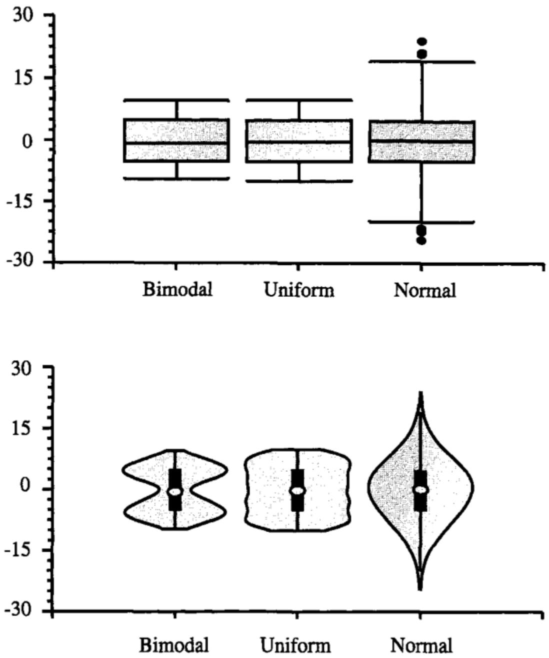

Figure 2.6: *The same three distributions (bimodal, uniform and normal) are represented by conventional box plots (top) and violin plots (bottom). The violin plot does a better job in conveying information about the underlying distribution. [26]*

Another extension to the traditional box plot, called *violin plot*, was proposed by Hintze and Nelson [26]. In its core it features just a visual modification of a box plot. The box is drawn in solid black, while the median is represented by a white dot within the box. This is supposed to support a quick comparison of medians of multiple juxtaposed plots. Around this core, a density plot is drawn symmetrically to the center axis to convey information about the underlying distribution. Figure 2.6 shows a comparison of violin plots and box plots representing the same data.

A similar visualization, which yet conveys even more information about a distribution, was proposed by Potter et al. [48]. The so-called *summary plot* visualizes the mean, kurtosis, first and second standard deviation, skew and tailing, as well as the density in addition to the information a traditional box plot presents. All these data elements are represented by icons placed along the center axis. The distribution density is visualized similarly as in the violin plot, but also adds a redundant color coding on top of the histogram. On top of the described visualization, which can be seen in Figure 2.7, the density of a well-known distribution is plotted in dotted lines. This is supposed to further help the user understand the underlying distribution, by comparing it to a well-known one. This reference distribution is either the automatically chosen, best fitting curve or a curve manually chosen by the user.

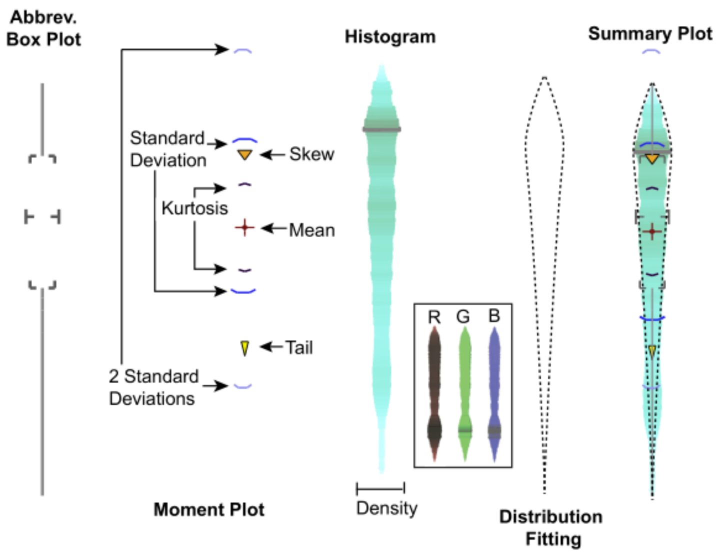

Figure 2.7: *A summary plot consists of an abbreviated box plot, various icons representing important moments, a histogram and a reference distribution. These elements are visualized over each other, which leads to a lot of information packed into relatively little screen space. [48]*

As can be seen in Figure 2.6, a major shortcoming of the traditional box plot is that fundamentally different distributions may lead to the same five summary values, which subsequently also lead to the same representation. This fact is further highlighted in Figure 2.8. Choonpradub and McNeil [12] proposed an adaption of the box plot that should solve this problem. In contrast to the methods presented previously, their goal was to not change or replace too much from the original representation. One reason for this is that these changes often introduce new problems, like the choice of the extent of smoothing. In their version of the box plot, which is shown in 2.9, the line thickness of the median, upper and lower quartiles are modified to indicate the distribution's skew. If the distribution is skewed toward lower values for instance, the lower quartile is emphasized. If the underlying distribution is bimodal, both quartiles are drawn thicker.

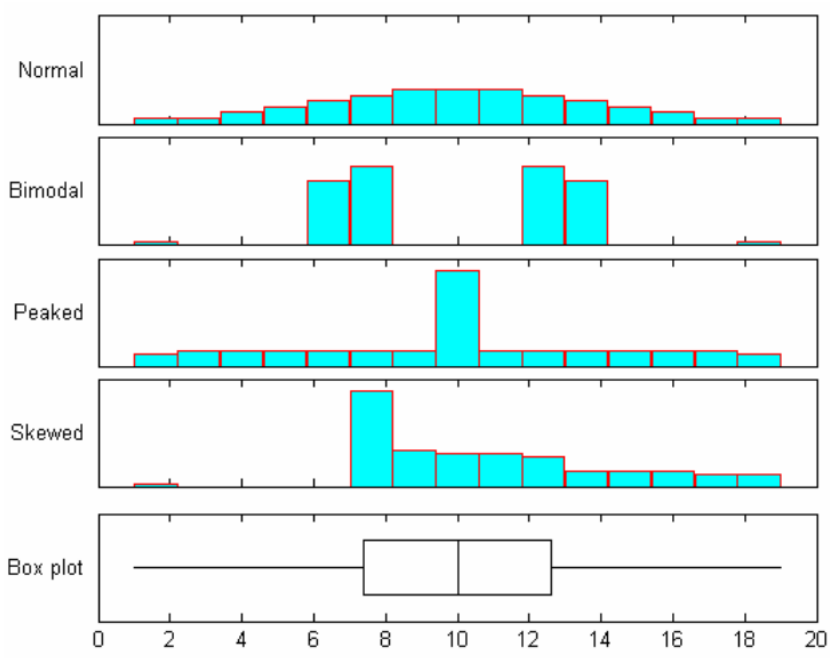

Figure 2.8: *The four presented distributions lead to the same values of median, lower and upper quartiles, minimum and maximum and therefore their box plot representation looks exactly the same. [12]*

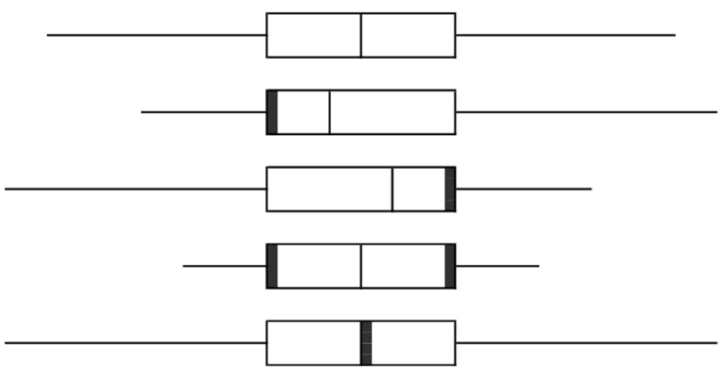

Figure 2.9: *The box plots by Choonpradub and McNeil [12] modify the thickness of the median, upper and lower quartile markers based on the underlying distribution. Distribution shown from top to bottom: normal, right-skewed, left skewed, bimodal, and centrally peaked. [12]*

There are many ways to visualize a distribution which can be used to represent the temporal uncertainty of events [58, 8, 26, 48, 12]. All of them rely on experience with statistical models of the user, to be used effectively [13]. This dependency of prior knowledge can be circumvented, if the presented information is framed in terms of discrete events and outcomes [25]. I.e. people can estimate probabilities more accurately, if the information is presented in countable outcomes, instead of a graph for instance. Based on this finding Kay et al. [31] developed the *quantile dot plot*. Instead of visualizing the density of a distribution in the traditional way, it is discretized into a fixed amount of dots. To determine the amount of dots at a certain value, the density function is evenly sampled. The sampling of a density function is illustrated in Figure 2.10 and Figure 2.11 shows some example dot plots.

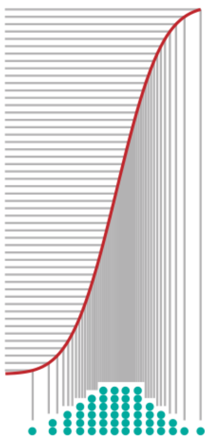

Figure 2.10: *The density function is uniformly sampled along the vertical axis. The samples are places within the interval [0,1] and their predefined number defines how many dots are used for the dot plot. The samples are gathered in bins along the horizontal axis. [31]*

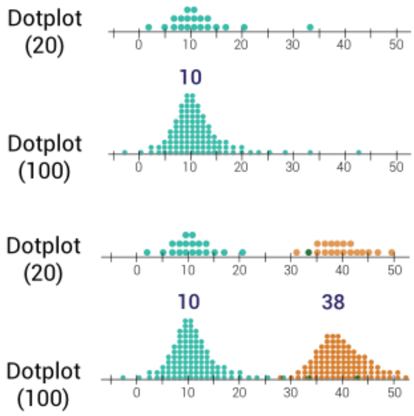

Figure 2.11: *These are example quantile dot plots with 20 or 100 samples respectively. [31]*

In some cases the exact distribution of temporal uncertainty is either not relevant or simply unknown. If the outer bounds of the uncertainty are known, we can speak of *bounded uncertainty*, in contrast to *statistical uncertainty*, which entails knowledge of the underlying distribution. Olston and Mackinlay [45] propose that these two kinds of uncertainty should be visualized differently, because of their fundamentally different properties. They argue that error bars are a good technique to convey statistical uncertainty, but should not be used for the first type. For the representation of bounded uncertainty, they propose a technique called *ambiguation*. It can be used to adapt traditional visualizations without uncertainty. To implement it the hard points, lines or surfaces that represent values, are drawn in a lighter color at places where they are uncertain. The uncertainty is bounded by the extent of the light color, while the default color of the graphic represents certain values. Figure 2.12 illustrates this principle for traditional bar charts, scatterplots, line charts, stacked bar charts and pie charts. Because of the adaptability of this technique, it can easily be used to modify a traditional visualization of temporal data to also incorporate uncertainty.

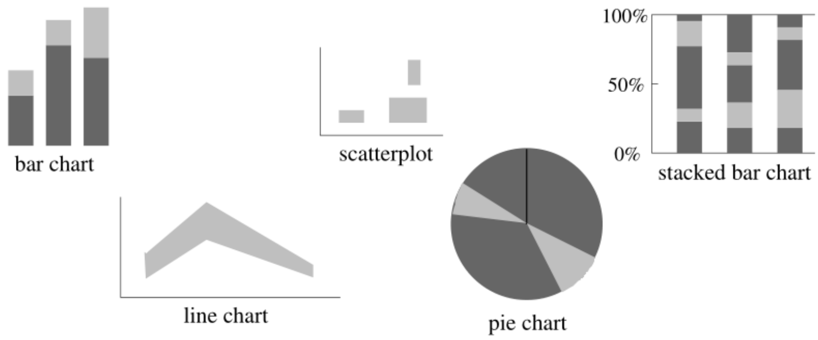

Figure 2.12: *The ambiguation technique shows the uncertain intervals of values in a lighter color. bar charts: The lower bound of every bar is represented by a traditional bar. On top of it is the uncertainty interval in lighter color. line charts: Instead of one solid line a surface is drawn in a lighter color, which encompasses every possible line. scatterplots: Instead of solid points there is a surface, in which a point may lie. pie charts: The lighter areas show where a borderline between two wedged may be. stacked bar charts: Just like with pie charts, the possible locations of borderlines are represented in a lighter color surface. [45]*

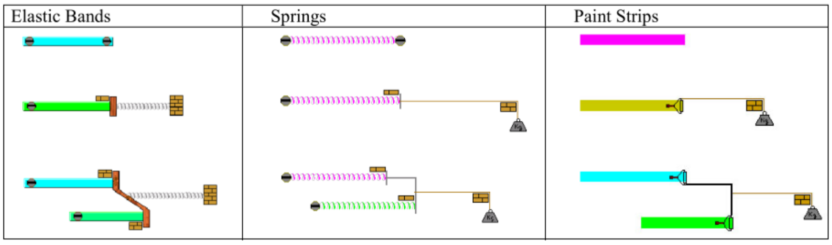

Figure 2.13: *The elastic band visualization suggests that the spring can pull and bend the elastic bands to an uncertain distance. This distance is bounded by walls. In the spring visualization the springs are bent by weights. Paint strips are also extended to an uncertain amount, because weights pull on their paint rollers. [11]*

Chittaro and Combi [11] propose three techniques that are similar to ambiguation and are also fit to visualize bounded uncertainty. The biggest difference is that these approaches are heavily focused on metaphors to promote intuitiveness. The three visualizations are called *elastic bands*, *springs* and *paint strips* and work in a very similar fashion. The only difference is their respective metaphor. They are all composed of a solid colored bar that represents the certain part of an interval. In the case of elastic bands and springs, certain boundaries are marked by a screw icon to symbolize that they are fixed at this point and cannot be moved. Bars in the paint strips visualization are not 'screwed on' because they represent paint on a wall, which cannot be moved anyway. In all three approaches uncertainty is expressed through a moving mass system, which can physically bend the elastic bands or springs or move paint rollers to extend the length of the bar. The uncertainty is always bounded by some obstacle metaphor, which prevents the springs and bands to extend farther and stops the paint rollers from moving beyond those bounds. Examples of all three visualizations can be seen in Figure 2.13. Multiple dependent uncertainty intervals can be linked by a solid connection to represent a fixed relationship of them.

The visualization of temporal uncertainty can often be important for the planning and management of complex projects. Such projects are often split into smaller tasks. The time it takes to fulfill these tasks can often be only roughly estimated. Furthermore, these tasks can be closely dependent on each other, which means that a delay of one of them leads to a delay of the whole project. To visualize these dependencies the US Navy developed the *Program Evolution and Review Technique* (PERT) charts. These charts depict every task as boxes which are connected by arrows to symbolize inter-task dependencies. The time it takes to fulfill a task, as well as the corresponding uncertainty is only given in numerical values and not visualized explicitly.

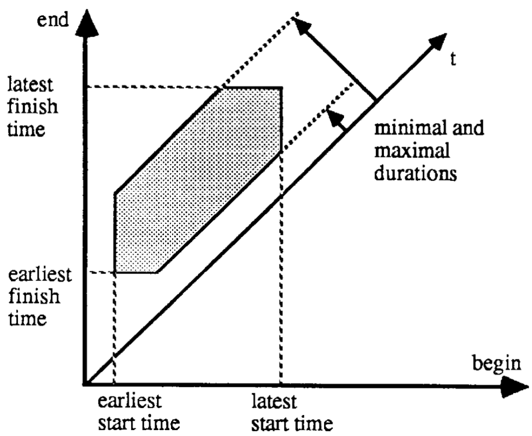

Figure 2.14: *The horizontal axis is used for possible starting times, while the vertical axis represents possible finish times. The earliest starting time and finishing time, as well as the latest starting and finishing time respectively, form opposing corners of a rectangle. This rectangle is further constrained by cuts parallel to the 1st median line, which represent the maximal and minimal durations. [49]*

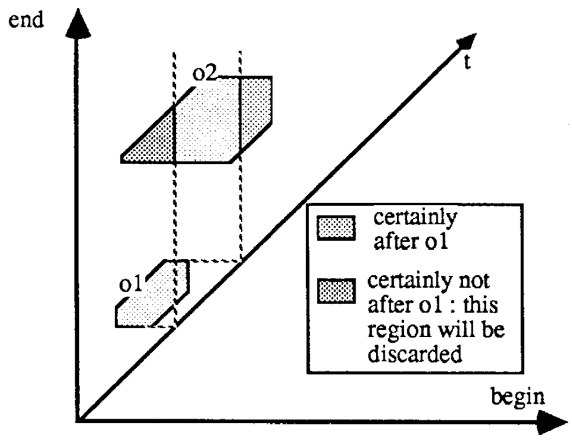

Figure 2.15: *Two SOPOs are compared by propagating their constraints onto one another. In this example the earliest and latest finish times of o1 are propagated to o2, to find out in which cases o2 will take place before o1 has finished. [49]*

In 1896 Rit [49] proposed a technique that visualizes the temporal bounds and all possible intervals of an event with uncertain start and end times. The technique, called *sets of possible* *occurrences* (SOPOs), was utilized in a system for medical planning [41]. In this approach the time domain is split up into two axes. The horizontal axis depicts the interval of possible start times, while the vertical axis encodes all possible end times. A resulting SOPO, defined by a starting interval, an end interval and its maximum and minimum duration, can be seen in Figure 2.14. It is a surface which encompasses all possible time intervals in which the event can take place. These intervals are represented by points within the diagram. Multiple SOPOs in the same view can also be compared, by propagating their associated constraints onto one another, which is demonstrated in Figure 2.15. As already mentioned before, SOPOs were integrated into a medical planning system, called SOPOView [41] and evaluated in practice. The field experiment showed that most physicians who worked with the system, found the visualization too complex and confusing. It took them too long to familiarize themselves with the representation to be able to use it effectively.

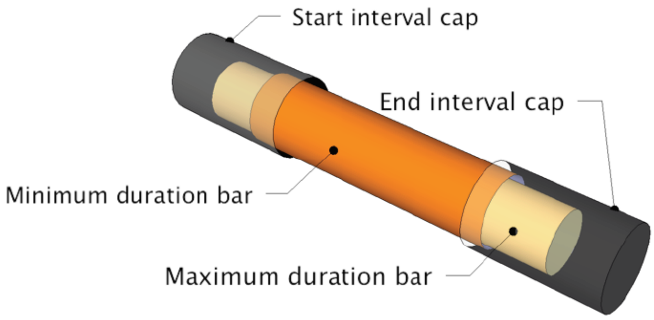

Figure 2.16: *The two inner cylinders encode the minimum and maximum duration respectively, while their possible extent is conveyed through two bounding caps. [4]*

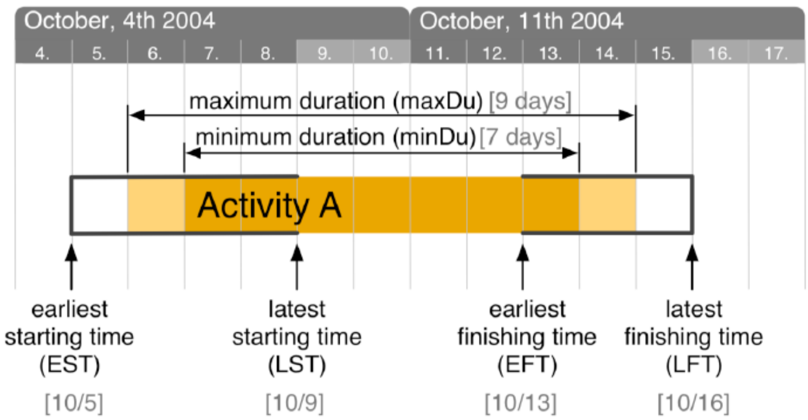

Figure 2.17: *The 2D version of the visualization works analogous to the 3D glyph. The inner bars encode the minimum and maximum duration, while the two caps represent the event's bounds. [4]*

There are further approaches that try to incorporate temporal uncertainty into planning systems. One technique was proposed by Aigner et al. [4] and is called *PlanningLines*. Their visualization consists of two black boundaries that represent the earliest and latest start and finish times. Furthermore, there is a centered bar, which represents the minimum duration of the event, while a the maximum duration is marked by a lighter color extending the first bar. These bars can metaphorically be moved within their bounds to make up all possible event occurrences. Figure 2.16 shows a 3D glyph version of the PlanningLines visualization and Figure 2.17 shows a 2D version. Figure 2.18 shows how the approach is incorporated into a simple project plan.

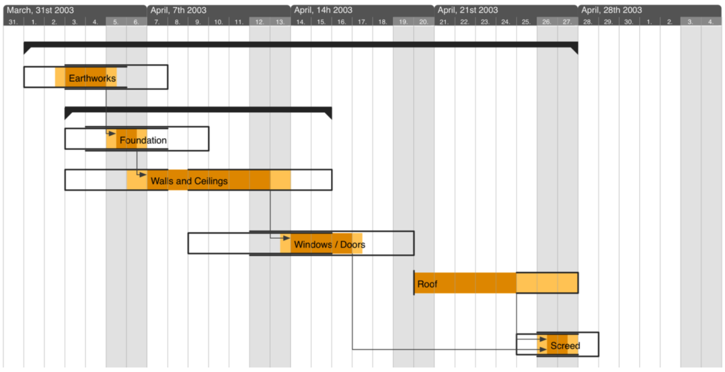

Figure 2.18: *This project plan utilizes Aigner et al. [4]'s visualization to show temporally uncertain events and their dependencies. Dependent events are connected through arrows. The possible extent of the whole project is visualized through two additional black bars. [4]*

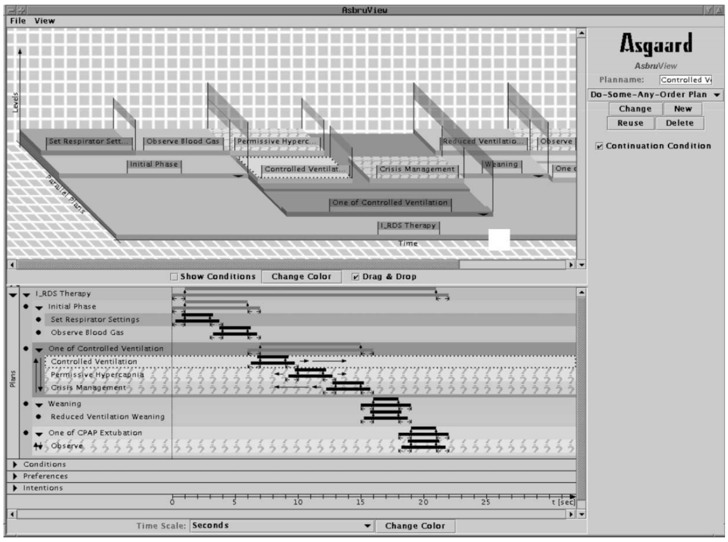

Figure 2.19: *The main interface of AsbruView shows a topological view at the top and a temporal view in the bottom. The temporal view utilizes Time Annotation Glyphs to represent events. [34]*

Another planning system, called *AsbruView*, was developed by Kosara and Miksch [34]. The main interface of AsbruView is shown in Figure 2.19. Its purpose is to help physicians in planning therapy procedures for their patients. Traditionally, clinical protocols and guidelines are represented using flow-charts, decision tables or plain text documents, which are not well suited to display complex medical procedures. A far better suited alternative would be the use of *Asbru*, which is a plan representation language. The problem is that Asbru requires expert knowledge to be used and is therefore not suitable for physicians. AsbruView aims to close this gap between applicability and usability, by providing an intuitive user interface to the underlying planning language. To represent the plans, a visualization called Time Annotation Glyph was developed which can also incorporate the temporal uncertainty of the bounds of events. An event with uncertain start and finish times is defined by six values: the earliest and latest starting shift (ESS/LSS), the earliest and latest finishing shift (AFS/LFS) and the maximum and minimum duration (MaxDu/MinDu). Some of these values may also be undefined or uncertain. The resulting glyph can be seen in Figure 2.20. The length of the two bars encode the MinDu and MaxDu respectively, while the other temporal marks are represented by vertical lines and labels along the time axis.

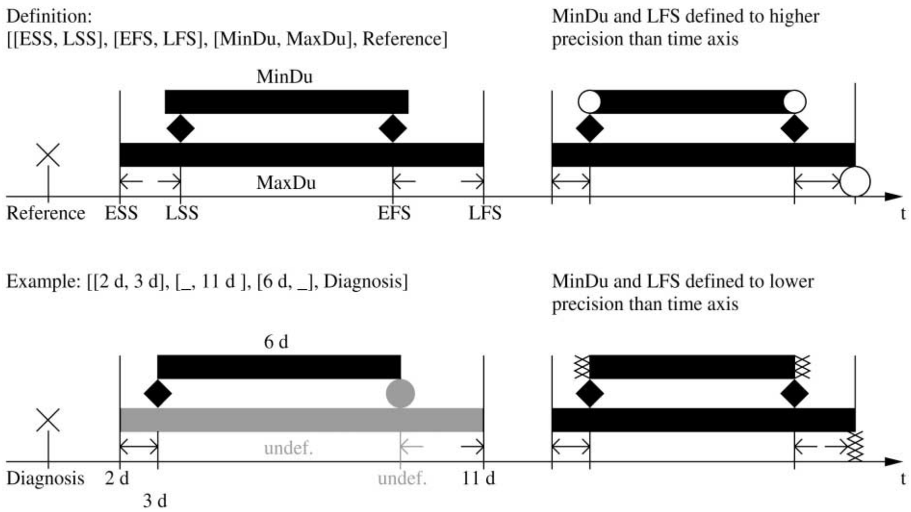

Figure 2.20: *Upper left: Shows a fully defined event visualized through a Time Annotation Glyph. Lower left: Shows an event with an undefined EFS and MaxDu. Upper right: Shows that temporal marks can change their appearance to white circles if the definition granularity of the marks is higher than the temporal granularity of the plan. Lower right: Shows that uncertainty of temporal marks is represented by zig-zag patterns. [34]*

All the approaches presented in this section so far are concerned with the temporal uncertainty of an arbitrary event that does not necessarily have to be defined in any other way (0D case). It would also be possible to show a temporal uncertainty along a linear path of some kind. An example would be the arrival time of a bus along its route. This would mean that the uncertainty does not only have to be visualized at one point but along an axis (1D case). This dimensionality can obviously be increased indefinitely, but especially the next logical step (the 2D case) is of great importance to many researchers. Being able to visualize Probability Density Functions (PDFs) over a 2D field (e.g. a map) is of great interest to Global Information System (GIS) specialists.

There have been many approaches to solve this problem, especially from the GIS community [47, 29, 17, 39, 53, 46]. An interesting example is the system of Potter et al. [47], which compares the PDFs defined over a 2D field with a reference distribution and visualizes the resulting difference metric via a color code over the field. An overview of this system can be seen in Figure 2.21. Another approach that aims to visualize uncertainty in spatio-temporal data was proposed by Shrestha et al. [53]. The so-called *Storygraph* consists of two vertical axes to encode two spatial dimensions. A spatial position is given by a line between those two axes. The time is encoded as a horizontal axis in between. An event is therefore represented by a point along a line. This principle is illustrated in Figure 2.22. To also incorporate spatial and temporal uncertainty, instead of fixed points and lines, uncertainty intervals are used. This extension is very similar to the ambiguation approach proposed by Olston and Mackinlay [45]. Two examples for the visualization of uncertain data with this technique are presented in Figure 2.23.

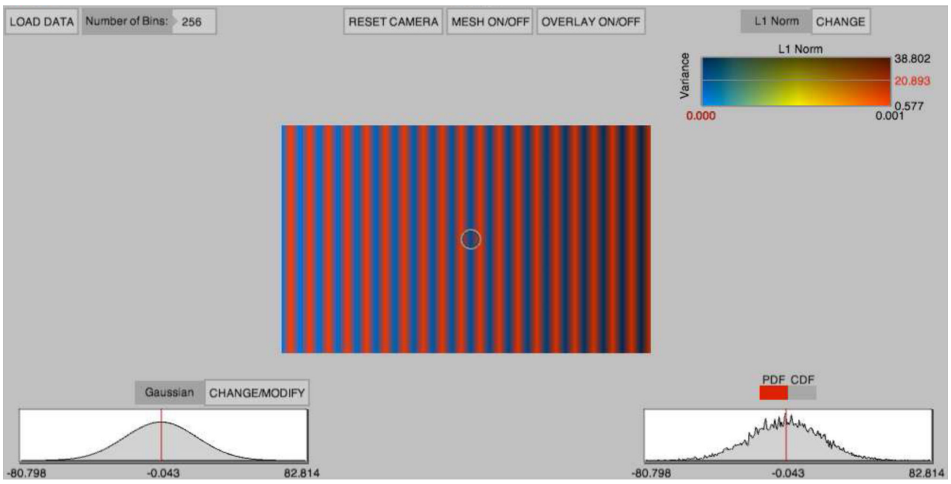

Figure 2.21: *In the center of the interface the 2D field is presented. It is color coded based on the difference metric of the respective PDF at every point compared to the reference distribution. These distributions are also respectively visualized in the bottom left and right corners. In the top right corner the legend of the color code is given. [47]*

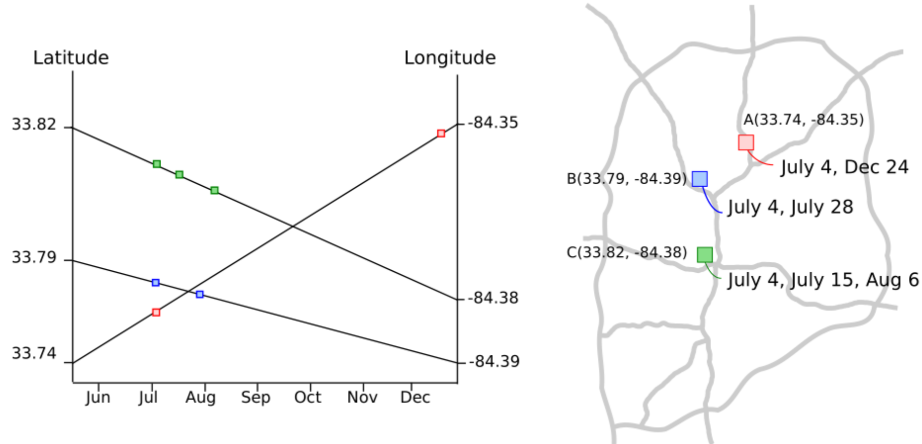

Figure 2.22: *The Storygraph on the left visualizes the same events that are listed on the conventional map on the right. By depicting a linear time axis, temporal relationships between multiple events are easy to detect. [53]*

The visualization of temporal uncertainty embedded into a spacial domain is certainly very interesting and important in fields like GIS science, but it is not the main focus of this work. The presented overview and few examples given should provide the interested reader with a starting point for further investigation. The rest of this thesis is focused solely on the 0D case of temporal uncertainty. This means that the depicted events are not mandatorily spatially bound, or rather their spatial occurrence is not explicitly depicted in the utilized visualization approaches.

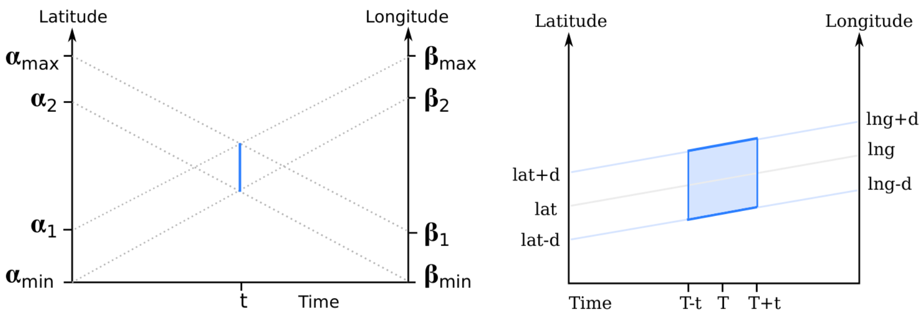

Figure 2.23: *The Storygraph on the left shows an event with spatial uncertainty. Instead of two certain points along the spatial axis, there are two intervals. Therefore, the point that usually represents an event becomes a vertical line. The same applies to the right story graph, which also features temporal uncertainty. The result is a surface instead of a point representation of the event. [53]*

One way of looking at the state-of-the-art of temporal uncertainty visualization is to find out which possible tasks there are existing methods for and which still lack a fitting solution. To categorize all the different scenarios for temporal visualizations a taxonomy is needed. A suitable taxonomy for the given topic is given by Aigner et al. [3], who propose a systematic view on the visualization of time-oriented data. Their approach to categorizing techniques is based on three practical questions:

1. What are the characteristics of the time?
2. What is analyzed?
3. How is it represented?

The first question further splits up into sub-questions. The first determines if there are time points to be visualized or time intervals. The second one regards the structure of time, which can be linear, cyclic or branching.

The second question splits up into three sub-question, which further leads to two categories each. The first one asks about the frame of reference, which can either be abstract or spatial. The second one determines if the number of variables is either univariate or multivariate. The third distinction regards the level of abstraction of the visualized data, which can be either raw data or a data abstraction.

The last question splits into two sub-questions, which form two distinctions each. The time dependency can either be static or dynamic and the dimensionality of the visualization can either be 2D or 3D.

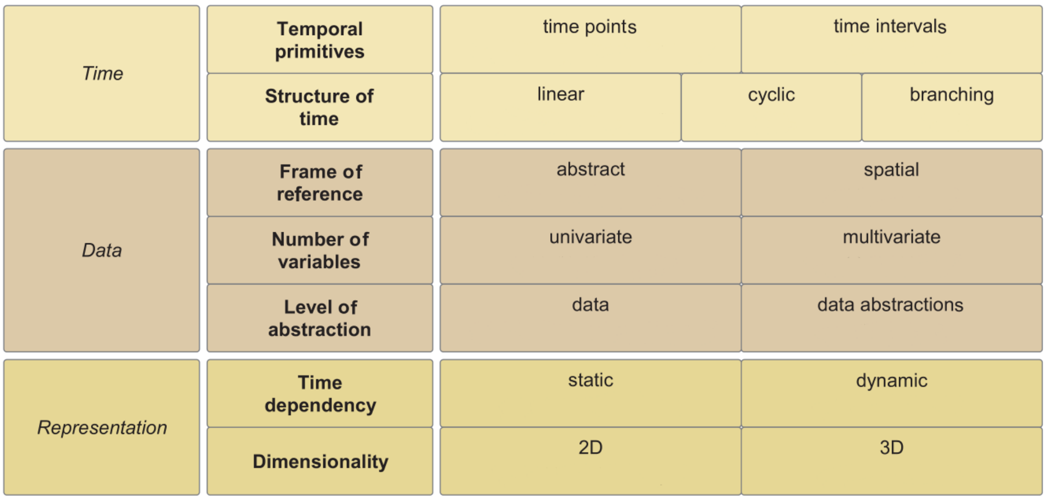

Figure 2.24: *This illustration shows how the initial three questions on the left are further split up into sub-questions, which each split up all techniques into two or three categories each. [3]*

All these questions lead to the categorization illustrated in Figure 2.24. If a visualization is categorized, every question has to be answered, which leads to a specific category. This means that a specific category is defined by the combination of answers to each question. Hence, the taxonomy leads to a number of 192 distinct categories.

Originally this taxonomy is meant to categorize visualization techniques for time-oriented data, which do not necessarily incorporate uncertainty. But since temporal uncertainty only means an additional type of data added to the time-oriented data, the same taxonomy can be used to categorize techniques for the visualization of temporal uncertainty. The challenge is that 192 categories are a vast number to find approaches dedicated to temporal uncertainty for. It is also apparent that there are sub-categories which hold hardly any such approaches. For instance, there are hardly any approaches that deal with branching time and temporal uncertainty at the same time.

The gist of this is that given a specific scenario, by answering all questions of the categorization, it is probable that there is no existing technique for this category, which also incorporates temporal uncertainty. Obviously this does not mean, though, that the given data can not be visualized effectively. To tackle this problem Brodlie et al. [9] compiled a catalog of visualization techniques, which allow traditional techniques to be extended to also encompass temporal uncertainty. This way visualization designers can pick the right approach from all visualizations fit to represent time-oriented data and extend them to also encode uncertainty.

2.3 User study design theory

To improve the design of the practical studies of this thesis, literature about the theory of study design was reviewed. Especially works which feature common pitfalls and valuable advice are most important to us. In this section all the relevant literature that has been found is presented.

Kosara et al. [33] states that visualization is mostly practiced as a craft and that techniques might sometimes be well motivated by theory, but still do not work well in practice. Therefore, it is not enough to justify a technique by theory, but its performance must also be measured. A good way to gauge this performance is through user studies. The issue is that user studies have to be well designed to yield valuable results. Furthermore, a user study needs to be the right approach for a given problem or to answer a given question in the first place. Kosara puts great emphasis on this initial decision, whether a user study is the right choice or an alternative method should be used. To design, implement, run, and analyze a user study properly is very time consuming, which makes this decision a very important one. Generally, this effort is worth it if the goal is to answer a relatively small question, like whether one technique outperforms another one in solving a certain task. To answer bigger questions this way, some amount of generalizations are needed, which can be tough to get right and valid. Kosara also elaborates on different kinds of failed studies. Study results might be null results, inconclusive or not compelling. These three cases need to be handled differently and do not automatically mean that the study was worthless nor that the results should be discarded.

Tory and Möller [56] also feature a section about the methodology of user studies. Just like Kosara et al. [33], they state that user studies are generally expensive and time consuming. This is especially the case if the study is aimed at evaluating a whole visualization system, since this makes the design of the study very complex and the needed extent of it, to cover the whole system, also grows quickly. For this reason it is usually better to focus on smaller questions and only evaluate single tasks or visualization techniques through user studies at a time.

The work of Hullman [27] is especially interesting in regard to this thesis, since it is focused on evaluating visualizations concerned with uncertainty. Furthermore, this work does not only fit thematically very well, it also gives clear guidelines for the design of user studies. These guidelines are split up into three categories, concerned with different issues. The first category of suggestions tackles the problem that the subjective view and understanding of uncertainty of a study participant might influence the results. The understanding, for instance, can be checked throughout an experiment, by checking if the given answers of a participant are logically sound (e.g. estimated probability might need to add up to 1 at some point). The second category of guidelines is concerned how answers are elicited from participants, without impacting the results with the elicitation itself. The best way to ask a certain question can, for instance, be tested in a pilot test. The last category sheds light onto the problem of heuristics used by the participants. These heuristics are shortcuts how one can arrive at a certain answer to a posed question. Through such a heuristic a participant might not even need to pay attention to a given graphic, which should be the focus of a study. For this reason it is important to think about possible heuristics during the design of a user study and also to look for signs of heuristics in responses when evaluating the results of a study.

In Kinkeldey et al. [32]'s work they take a closer look at the correct way of assessing visual representations of uncertainty. Because their main focus lies on geospatial data, they systematically review user studies in this domain. While the specific domain of this work does not match the one of this thesis, we still believe that the presented findings are also of value in the context of temporal uncertainty and its evaluation. One of their main contributions is the identification of five main goals of user studies and their corresponding measures. First and foremost there are experiments that measure *user performance* [22, 51, 14, 40]. This is usually done by analyzing the accuracy of given answers and the speed at which those answers are given. The second type of study tries to measure the *acceptance* of a new visualization. In the context of uncertainty visualization this could be experiments that try to find out if users prefer to have uncertainty explicitly visualized, or rather not be presented with this additional information. Existing studies in this direction yielded varied results. While some endeavors concluded that the additional information tends to clutter the view and had negative effects [52, 54, 30], other findings indicated uncertainty information can clarify the view [1, 5, 15, 35, 37, 59]. The conclusion Kinkeldey et al. [32] draw from this is that appropriate solutions need to be found to visualize uncertainty in a way that is helpful and acceptable to the user. Another measure that can be taken into account is the *user confidence* while giving answers during a study. This measure is often used additionally to performance evaluations. While this information can be valuable to find lucky hits within the collected answers (i.e. a user could not determine the right answer, but still gave it by guessing correctly), the further value of it seems questionable, since it usually correlates strongly with the user's performance. A similar measure is the *user preference*, which asks study participants for their personal opinion. Interestingly, this measure does not necessarily correlate directly with the performance measure. This means that it may be important to additionally ask for user preference if the evaluated visualization should be deployed in a real scenario. The fifth measure that is identified is the *intuitiveness* of a representation, which most user studies do not concern themselves with. We believe this measure to be very important, especially if the evaluated visualization is not only supposed to be used by experts, but also by non-expert users. Regarding the general approach of user studies Kinkeldey et al. [32] state that there is no methodology commonly agreed upon. This is not only true for the domain of geospatial visualization or uncertainty visualization, but reaches much farther. This is also a reason why it is important to review existing approaches and learn from them, until a systematized approach for empirical studies is proposed and widely adopted in the InfoVis domain. In regard to the evaluation of user studies, two main goals are identified. The first goal is the improvement of a visualization by comparing multiple alternatives and see which one performs 'better' (in regard to the measures explained beforehand). The second goal is to gain deeper insights into cognitive processes of the study participants and understand why one representation works better than another.

Lam et al. [36] did an extensive literature review of about 850 publications with the goal of giving a novel approach to determining the most effective evaluation of a given visualization. Their approach works through defining seven scenarios with their respective goals/outputs, typical evaluation questions and fitting methods and stands in contrast to existing approaches, which are usually based on listing existing methods. Since there is a many-to-many mapping between scenarios and methods, this new approach is supposed to limit the development of new methods less than conventional approaches. The following scenarios are defined: *1. understanding environments and work practices (UWP), 2. evaluating visual data analysis and reasoning (VDAR), 3. evaluating communication through visualization (CTV), 4. evaluating collaborative data analysis (CDA), 5. evaluating user performance (UP), 6. evaluating user experience (UE) and 7. evaluating visualization algorithms (VA).* These scenarios can be categorized into two main groups. The ones that evaluate an underlying process and the role of a visualization (UWP, VDAR, CTV and CDA) and those that focus the evaluation on the visualization itself (UP, UE and VA). The literature review shows that a vast majority of 85 per cent of the evaluation done falls into the second category, with two thirds being evenly split up between evaluation of user performance (UP) and experience (UE). Since the study presented in this thesis also falls into the UE scenario, we take a closer look at it. While performance evaluation is focused on providing reproducible results, UE is about collecting subjective user reactions to inform the design of visualizations. In terms of observing users and collecting subjective information it can be similar to UWP, but its main focus is on the visualization and not the context it is situated in. Typical questions in UE are *'What features are seen as useful?'* or *'Is the tool understandable and can it be learned?'*, which are of similar nature to our own research questions. There are three typical methods listed for UE. First, there is the *informal evaluation*, in which domain experts test a given system informally and give their feedback. The second listed method is the *usability test*, in which the participants are observed during the solution of predefined tasks. Lastly, there is the *laboratory questionnaire*, which comes closest to describing the approach of this thesis. It is often done additionally to an UP, by asking for subjective information, for instance through Likert Scales.

2.4 User studies in InfoVis

To optimize the design of a user study it is helpful to not only look at the theory behind it, but also survey existing studies that have been put into practice before. By reviewing these experiments we learn more about the state-of-the-art of study design and evaluation. In this section a selection of user studies similar to the practical part of this work are presented and reviewed.

The user study that guides this thesis the most is the one by Gschwandtner et al. [22], as this work builds up on theirs. This study compares six different techniques for the visualization of uncertainty in the temporal domain. To determine which technique works best for certain tasks, five different types of tasks were designed. The first type is about finding out how users interpret the different visualization techniques. In the second type of tasks the users are asked to read the boundaries of uncertainty intervals from the visualization. The third type is about determining the extent of an uncertainty interval. In the fourth type of tasks, the users have to gauge certain probabilities using the visualization and the last type of tasks asks the users for their opinion about the visualization.

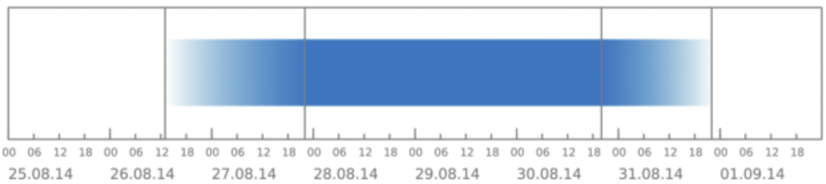

Figure 2.25: *A gradient plot shows the certain parts of an interval as a solid color, while the uncertain parts are represented by a color gradient. [22]*

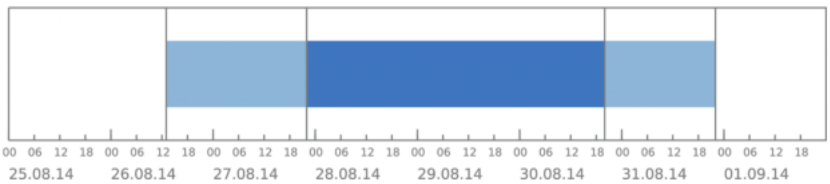

Figure 2.26: *This technique, called ambiguation, shows the uncertain areas of an interval in a lighter color than its certain part. [22]*

The actual study was conducted with 73 participants who were all bachelor students in computer science. The students were recruited from a course in information design and visualization, which implies a certain knowledge about this topic. To automatically track relevant data, such as completion time and accuracy during the study sessions, the EvalBench software library [2] was utilized. This library was designed especially for the evaluation of visualization. To analyze the results Gschwandtner et al. ran an ANalysis Of VAriance (ANOVA) for each task and subtask and backed up their results with a non-parametric Kruskal-Wallis test. Their analysis showed that the technique *ambiguation*, which can be seen in Figure 2.26 and is also reviewed in Section 2.2, works best for tasks in which the user has to judge the exact duration and bounds of an uncertainty interval. If the user has to determine certain probabilities within the uncertainty interval, *gradient plots* (see Figure 2.25) work best.

Kay et al. [31] based their work on an existing smartphone application, called *OneBusAway* [16]. In an initial study they asked users about their opinion regarding the most important requirements of a public transit application. To elicit those answers, every participant was provided with suggested questions, which they had to rate based on how often they try to answer those questions using *OneBusAway*. Furthermore, the participants were questioned about additional information which the application does not provide and its potential helpfulness. Through this survey the most important design requirements and a detailed description of user needs could be identified.

Based on these results Kay et al. designed a visualization for bus arrival time predictions aimed to be displayed on small screens. In their final design, which was developed in an iterative process, they introduce *quantile dotplots*, which are modified dotplots that are a discrete analogue to common probability density plots. An example plot can be seen in Figure 2.10. To evaluate their novel design, a large user study was conducted. In total, 441 participants were presented with different visualizations, which can be seen in Figure 2.27, and asked to solve tasks in which they had to estimate different probabilities. When evaluating the results Kay et al. [31] were more interested in the answers' variance than in the answers' overall accuracy. The reason for this is that very consistent but skewed estimates can easily be compensated by experience if the user utilizes an application over a longer period of time, while inconsistent estimates cannot be dealt with that easily and therefore indicate bad visualization design. The results of the study show that the novel quantile dotplots perform better than common density plots in terms of variance, while they are not as visually pleasing.

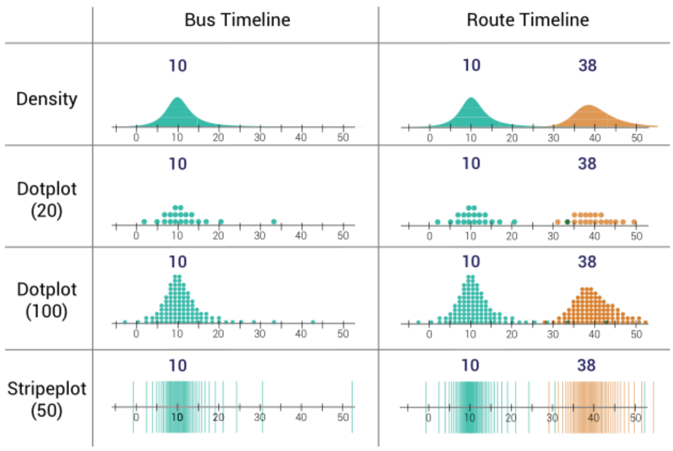

Figure 2.27: *Shows the four different visualizations the participants of Kay et al.'s study were presented with: a common density plot, a coarse and a finer dot plot and a stripe plot (top to bottom). [31]*

The main goal of this thesis is to find out more about the intuitiveness of representations of temporal uncertainty. These insights are valuable for the design of new visualizations, especially those aimed at non-expert users. MacEachren et al. [40] also tried to find out more about the intuitive design of visualizations through user studies. To do this they conducted two separate user studies, of which the first one shares a similar goal to the work of this thesis. Their first study compares many different sets of symbols for the visualization of uncertainty, to find out which are most intuitive to people. Every set consists of three symbols, which encode high, medium and low uncertainty of 3 different kinds (accuracy, precision and trustworthiness), for 3 data domains (spatial, temporal and attribute). Some example sets can be seen in Figure 2.28. In total 102 symbol sets were rated by 31 undergraduate students for their intuitiveness on a scale from 1 to 7. After this first series of tests, the most unintuitive symbol sets were filtered out, which left 76 sets. Those were again rated by 72 participants with a background in GIScience.

Figure 2.28: *Every column shows a set of three icons which represent high, medium or low uncertainty respectively. [40]*

After this first study about intuitiveness, the 20 highest rated symbols for every combination of uncertainty type and data domain were compared in a second experiment. The goal of this subsequent study is to compare the selected visualizations' performance, so the combined results of both studies yield the best visualizations for a given task, which is intuitive and efficient at the same time. To compare the chosen symbols, two quadratic matrices with 9 symbols each were visualized side by side. The participants were asked to answer the question which of the two matrices featured a lower overall certainty, based on the presented symbols.

Walny et al. [60] conducted a study with the goal of providing deeper insights into the way people think about and use visualizations to communicate their ideas. A total number of 69 researchers were observed using whiteboards during brainstorming, thinking, communication and similar actions. Whiteboards were chosen as a visualization medium, because they support a variety of thinking tasks, like personal and collaborative cognition, group meetings and planning. The results of the study feature interesting insights, such as different uses of emphasis techniques and the usage of ellipses as a focus and context technique. The pre-study presented in this thesis aims to provide similar insights through a similar approach, by also observing users in their creation of visualizations and reviewing those drawings.

In another study of greater exploratory nature, Walny et al. [61] asked 22 participants (mostly computer science students) to sketch visualizations of a given dataset. The data was provided in a table format and was about appropriateness ratings of certain behavior in given situations. The students' task was to create visualizations to find interesting patterns in the data and articulate them in a post-sketching questionnaire. The results were analyzed through multiple coding passes, which showed that, even though 9 out of 22 participants claimed to have no experience in visualization, most of the sketched representations could be classified as known types. As already stated, the study is of exploratory nature and therefore it does not answer many questions, but rather raises interesting questions and gives direction to future research.

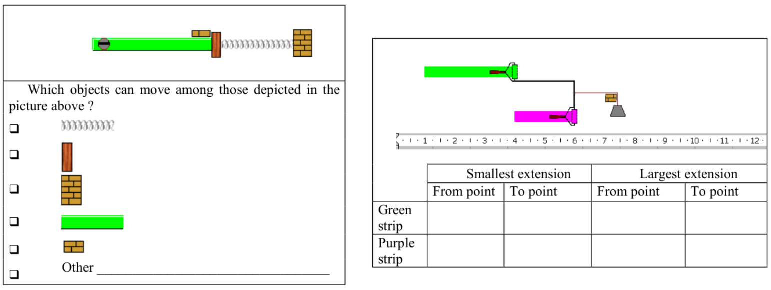

Figure 2.29: *On the left an example question from the first part of the first study by Chittaro and Combi [11] is shown. The user's task is to mark all the elements which seem to be movable in the given visualization. On the right an example question from the second part is shown. Here the user is asked to determine the bounds of the depicted uncertainty intervals. [11]*

In Section 2.2 of this chapter three techniques proposed by Chittaro and Combi [11] are presented. The same researchers evaluated their designs in two comparative user studies. The first study is further split into two main parts and aims to find out which of the proposed techniques is the most correctly understood. The first part consists of a single exercise, in which the participant is asked to identify the movable objects of the given visualization. This question is asked once for each of the three proposals in a randomized order. An example of this exercise can be seen on the left side of Figure 2.29. The second part consists of three exercises, in which the user is asked to determine the temporal bounds of uncertainty of a given visualization. An example of such an exercise can be seen on the right side of Figure 2.29. These three exercises are repeated once for each technique, with a randomized order of techniques. In this first study 30 participants (13 female, 17 male) took part, with ages ranging from 24 to 37 years. The results, which are presented in Figure 2.30, show that elastic springs and paint strips are significantly easy to be understood correctly. For this reason these two techniques were chosen to be compared in the second study.

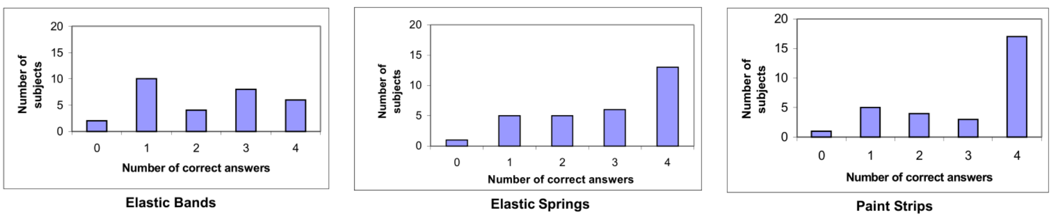

Figure 2.30: *The three charts show the number of correctly answered questions from the first study of Chittaro and Combi [11]. Since elastic bands performed significantly worse than the other two techniques, it was omitted from the following study. [11]*

The second study is split into two parts. In the first part the participants are asked to interpret given visualizations. This is done by presenting a visualization and offering three possible interpretations to choose from, of which one is correct. The second part of the study is about the generation of visualizations. The participants are provided with a tool that allows them to create elastic spring and paint strip visualizations themselves. After a training phase in which the participants have time to get familiar with the provided interface, they are asked to create visualizations satisfying given textual specifications. Since one study session consisting of all exercises concerned with one of the two proposed techniques took around 50 minutes, the two sessions were split up into two separate days, to prevent fatigue from affecting the results. The study was conducted with 31 participants (16 female, 15 male) with a medical background, between the ages of 23 to 44. To analyze the results the number of correct answers from the first part of the study and the number of completely correctly generated visualizations from the second part were counted. These sums were statistically evaluated with a Wilcoxon test for two related samples. Paint strips scored better than elastic springs on almost all accounts, but only by small margins and without statistical significants.

2.5 Discussion

In the preceding sections of this chapter a multitude of works of three main categories are presented. All these works are related to the practical part of this thesis and form the theoretical basis for the study design of this work. The most important points (i.e. the ones that guide our design the most) are compiled into a list and discussed. Every one of those points is again addressed in the following section to explain how the insights gained are incorporated into the resulting study design.

- **D1:** MacEachren et al. [40] try to find out which symbols of their test set are most intuitive for the visualization of certain types of uncertainty. To find the most intuitive ones, they directly let the study participants rate every symbol set's intuitiveness. This direct method stands in contrast to other methods which try to gauge intuitiveness by measures like completion time and/or correctness of task solutions. Both approaches have their merits, since the opinion of study participants is not necessarily reliant information. On the other hand, judging intuitiveness in an indirect way means to generalize study results, which may be tough to get right.

- **D2:** Since MacEachren et al. [40] work with a high number of symbol sets which are rated for intuitiveness, they utilize a pre-study to filter out the most unintuitive sets and conduct their main study with only the reduced set. This way it is possible to consider more possibilities of different visualizations and still gather enough data of the most interesting ones to draw statistical conclusions.

- **D3:** In the study of Walny et al. [60] visualizations drawn on whiteboards are analyzed. Walny et al. give several reasons why they believe that whiteboards are a good medium for spontaneous visualizations. They state four main characteristics which make whiteboards particularly suitable in this context. These characteristics are listed as *immediacy*, *messiness*, *sketchiness* and *forgiveness*. We believe that their choice of medium is well justified by their reasoning.

- **D4:** After Walny et al. [60] analyzed visualizations for their study, they conducted follow-up interviews with some of the participants. Through those interviews they managed to find out more about the reason why the visualizations looked the way they did and what they were supposed to convey. We believe that interviews like this could be very important for some types of studies to get deeper insights into the way study participants think about visualizations in a user study.

- **D5:** In both projects by Walny et al. [60, 61] which are presented in an earlier section, an open coding approach is used to categorize all the different hand-drawn visualizations they observed. This approach allows to find similarities between multiple drawings, without knowing what to look out for exactly beforehand. This way new and unexpected findings can be made. Open coding also is a way of handling the complexity that arises when a collection of hand-drawn visualizations needs to be analyzed.

- **D6:** In their exploratory study Walny et al. [61] do not only analyze visualizations based on their technique, but also take a look at how the drawings are bound to the provided data. Every visualization is assessed whether its nature is more 'numerical', which means that the provided data is directly represented, or if it is more 'abstract', which means that the data is manipulated and/or filtered before it is visualized. This additional categorization of sketches adds another dimension to the study results, which can be examined for interesting correlations to other result dimensions. The assessment of how 'numerical' or 'abstract' a visualization is, leads to a continuum into which every data representation can be put. Figure 2.31 shows what this continuum looks like.

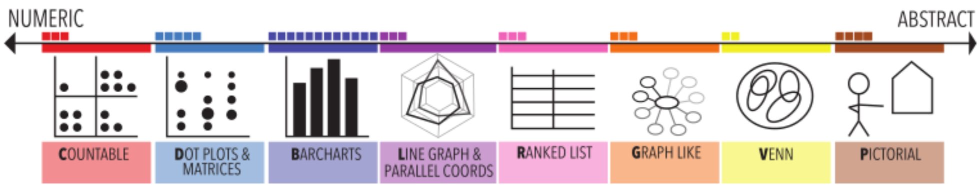

Figure 2.31: *The continuum of Walny et al. reaches from numeric to abstract and represents how much the underlying data was manipulated to reach its visual representation. A countable display, for instance, is directly bound to the underlying data, while a pictorial representation only visualizes some kind of abstract conclusion drawn from the data. [61]*

- **D7:** While the study participants of Walny et al.'s study [61] are kept anonymous, there is still some information given about them. The authors state how many male and female participants respectively take part in the study and also how many of them have a finished

Bachelors's, Master's or PhD degree. This information is not apparently used in the analysis, but we still believe that it is an interesting idea to assess this information and provide it, since it could be valuable for future follow-up research.

- **D8:** When it comes to comparative user studies, where multiple visual representations are tested against each other for different types of tasks, it is necessary to decide how visualizations and tasks should be distributed among the tested persons. For example, Robertson et al. [50] used a between-subjects design for their type of task (presentation or analysis) and a within-subjects design for the different visualizations and data sets.
- **D9:** Kosara et al. [33] point out that typically a user study is only suitable to answer small questions, while larger conclusions or generalizations based on the study results might not be valid. Hence, it needs to be considered whether a user study is sufficient for the examined research questions.
- **D10:** Furthermore, Kosara et al. [33] state that "null results" are a natural outcome of user studies, as the proposed hypothesis might not be supported by the data or the expected difference in results might not be significant. However, while results that are inconclusive or not compelling typically cannot be published alone, they still provide important information and insight for future research, and hence should not be considered a failure.
- **D11:** As explained by Hullman [27], it is crucial to ask participants in a way that provides all needed information and elicits an unbiased answer. Hence, questions need to be formulated carefully and in a direct way that allows for the most informative and insightful answers that can be compared to the ground truth.
- **D12:** Kinkeldey et al. [32] state that a general objective for every user study should be to recruit participants that represent the target user group as well as possible. However, most existing studies neglect this objective and simply recruit students, since they are easily available in a university context. These students are often chosen as representatives for the target group of expert users. Often they do have a sufficient level of theoretical expertise, but usually they lack practical work experience.
- **D13:** Another point advocated by Kinkeldey et al. [32] is that a training phase is important for effective empirical analysis. Only through an extensive training phase the scenarios, the data and the tasks can be sufficiently clarified. Furthermore, participants with different prior knowledge can be put onto the same level through training.

2.6 Conclusion

The previous section compiled a list of important points for the reviewed literature, which are relevant for the design of the user studies of this thesis (D1-D13). In this section every point is addressed again to explain how it is connected to the presented work and how the lessons learned are implemented into the study design (i.e. C1 addresses D1).

- **C1:** The Drawing Study is of exploratory nature. Its goal is to invent theories about the visualization of uncertainty. To reach this goal the participants of the study are asked to come up with their own ideas and concepts, so that hypotheses about the intuitiveness of visualizations can be derived. This means that the goal of our Drawing Study is different from that of MacEachren et al. [40] and therefore the approach is different too. The main study presented in this work, on the other hand, has a similar goal. It also compares the intuitiveness of different visualizations by directly asking the study participants to gauge the intuitiveness.
- **C2:** In MacEachren et al. [40]'s case the application of a pre-study makes sense because it enables a higher number of visualizations to be evaluated than without a preceding evaluation. In the case of this thesis, this approach is adopted, but the pre-study in form of our Drawing Study serves a different purpose. Instead of just filtering the visualizations to be tested, it is used to explore possible research questions to be tested in the User Survey.
- **C3:** Walny et al. [60] give several reasons why whiteboards are a good choice for the creation of spontaneous visualizations. Pencil and paper is the medium of choice for the Drawing Study. The reason for this is that it features all the desired characteristics of whiteboards described by Walny et al. [60] and in addition is easier to transport.
- **C4:** In the course of our Drawing Study it is paramount for us to understand in detail what our participants want to convey with their sketches. Since we are conducting the study separately with every participant it is not necessary to do follow-up interviews, as did Walny et al. [60]. The explanation of the drawing can happen during and right after the creation of the sketch, within each study session. In regard to our User Survey qualitative feedback is gathered through text fields accompanying each question which offer the possibility of further explanations.
- **C5:** The open coding approach of Walny et al. [60, 61] is interesting in regard to the Drawing Study, since there are also many hand-drawn sketches which need to be analyzed. This approach seems to be a viable solution to the problem of finding similarities between a great number of drawings.
- **C6:** Contrary to Walny et al. [61]'s analysis our Drawing Study is not analyzed in regard to the connection between the underlying data and the corresponding sketches. There does not seem to be an apparent benefit from arranging the sketches of this study along a continuum.
- **C7:** Walny et al. [61] gather some background information about their participants, which could be of interest when analyzing the data. In our case all participants of both evaluations are known to us. Therefore, additional information about our user group can be gathered without great effort. Furthermore, some personal information like the degree of education and the age of our participants is inquired at the beginning of our User Survey.
- **C8:** Since all questions asked during the main study concern the intuitiveness of the presented visualizations and do not pose a task to be fulfilled through its use, we do not

believe that there are important learning effects to be considered. For this reason a within-subject design will be adopted. However, to make sure that there are no learning effects or other effects caused by the sequence of questions, the order of questions will be different from one user to another.

- **C9:** Since user studies are only suitable to answer smaller, rather specific questions, our Drawing Study is conducted to find appropriate hypotheses to be examined. The resulting research questions are specific enough to be validly answered by the main study of this work.
- **C10:** In their work Kosara et al. [33] give suggestions about what is to be done with inconclusive study results. These suggestions can be valuable for any user study that fails to yield the expected results.
- **C11:** Especially our Drawing Study is vulnerable to suggestively asked questions. Special care needs to be taken to not influence the participants in their answers. Giving them any hints which visualization techniques might be applicable for the given data could have an unwanted impact on their sketches and drawings, which would render the study useless. The main study is more straightforward in this regard, since the posed questions are rather simple and can easily be asked in a non-suggestive way.
- **C12:** Since the target user group of this work is the general public, some students can conveniently be recruited as representative participants. However, there obviously must not be participants from university only. It is important to recruit people with various backgrounds and of different ages to declare it a representative sample.
- **C13:** During our User Survey it is important that every participant has the same understanding of the used visualizations and can interpret them correctly. To ensure that the survey starts with an introduction, which presents every used visualization with examples and a textual explanation.

CHAPTER 3

Drawing Study

The following chapter presents the so-called Drawing Study. It is split into three sections. In the first section, the introduction, the goal and rationale behind doing an exploratory study like this are presented. Furthermore, the basic approach regarding the study design and the reasons for it are explained.

The Design Section encompasses a detailed description of the concrete study design. It covers information about the selected participants, about the tasks and how the evaluation was conducted in practice.

The last section covers how the collected material of the study is analyzed to form hypotheses about the visualization of temporal uncertainty. Furthermore, a list of all generated hypotheses is presented.

3.1 Introduction

The main goal of this work is to generate insights about the intuitive visualization of temporal uncertainty. To find such insights an experiment in the form of a user study is a viable approach. However, as already mentioned in the last chapter's conclusion in **C9** and by Kosara et al. [33], user studies are only suitable to provide answers to smaller, more concrete research questions. This means that our goal needs to be defined more precisely before an appropriate experiment can be designed.

Our solution for this, as mentioned in **C2**, is similar to the approach of MacEachren et al. [40], who conducted a pre-study to filter out the most interesting visualizations from a predefined collection to test in their main study. Our approach differs from that, because we do not define possible visualization types in the first place. The reason for this is that we do not want to restrict the design space of possible visualization approaches, lest we might filter out interesting solutions. We utilize a pre-study with the goal of generating hypotheses toward the intuitive visualization of temporal uncertainty. These hypotheses are then formulated into concrete research questions for our main study.

Our goal is to gain insights about the intuitiveness of visualizations. We approach this problem by letting people design their own visualizations. The rationale behind this is that we believe that people utilize data representations and interactions which are familiar and understandable to them (i.e. intuitive to them). This means that by analyzing popular ways of visualizing temporal uncertainty by study participants, we can generate hypotheses about the intuitiveness of these approaches.

For this approach we want to restrict our participants as little as possible in their freedom of creating visual representations. This means that we do not want to use a computer supported tool that might have a learning-curve on its own and hence hinder our less skilled participants. For this reason, as mentioned in **C3**, we pick a similar approach as Walny et al. [60]. However, instead of using whiteboards, we opt to use a simple pen and paper setup. This option offers the same freedom to our participants to communicate their ideas freely. Furthermore, the conduct of the Drawing Study is easier, since a whiteboard would take up more space and cannot be used comfortably in many situations.

To make sure that every participant has the same understanding of temporal uncertainty and to give them direction in their creation of visualizations, we have to provide them with specific scenarios to visualize. According to the data-user-task triangle [42] we thereby provide the data for the visualization to be created. Furthermore, the target user group is provided by asking our participants to create something that is intuitive to themselves (i.e. each participants constitutes his/her own target user group). This means that only a specific task is missing to provide the basis for the creation of a visualization. Hence, every scenario also features a predefined task the visualization should support. The scenarios and tasks within these scenarios have to be defined in a way that ensure their generalizability. This way the results can be applied to many real world scenarios and tasks and are broadly applicable.

After collecting enough visualization designs from our participants the drawings have to be analyzed. As mentioned in **C5**, we tackle this analysis similarly to Walny et al. [61] with an open coding approach. This enables the identification of similarities between a large number of drawings, without advance definition of what to look out for concretely. These similarities can then be interpreted to form hypotheses about the intuitiveness of temporal uncertainty visualizations, which fulfills the goal of this preceding exploratory study.

3.2 Design

The Drawing Study is conducted separately with every participant. There is no uniform location specified for these separate sessions. The only important demand of the location is that it is a quiet enough place with few distractions to make focused work possible. Concretely we met with our participants at home (theirs or ours), at university or similar places. To enable our participants to express their visualization ideas visually and at the same time document their designs, pencils and sheets of paper were provided.

Every session starts with an oral introduction. At the beginning every participant is informed about the purpose of the study and its goal. Furthermore, the procedure is explained. The following points are also highlighted during the introduction:

- The sketched designs are supposed to be computer supported visualizations. This means that they may be interactive, may use animation and may have multiple colors. Since only a single pencil is provided, all these aspects may simply be visually indicated in the sketch and verbally explained.
- The assignment is to think of visualization designs and draw them as sketches. This does not mean that the sketching should restrict the visualization design in any way. If an idea is hard to draw, it should simply be drawn as well as possible and described.
- Since most participants are German native-speakers, we allow them to explain their designs in German. This way we prevent people from rejecting ideas because they are hard to explain. Only if they intend to use written words in their sketches, we ask them to write in English, so that the results can be presented more easily.

After all these points are made clear and potential questions have been answered the assignments are submitted to the participants. In total there are four assignments. Each consists of a scenario and a specific task to be supported by the visualization. According to the data-user-tasks design triangle proposed by Miksch and Aigner [42] this provided information is sufficient for the design of a suitable visualization. The target user group is the general public, which is represented by our participants themselves. The data to be visualized is given through the scenario and the task is provided explicitly. The same four scenarios are presented to each participant in the same order. We do not believe that there are any relevant learning effects involved. Furthermore, the third scenario builds up onto the preceding one, which makes their order necessary. The concrete scenarios are presented in detail in the next subsection.

Each scenario has an exact predefined textual description. This helps to keep the assignment and provided information uniform between each participant. At the same time it is paramount that every participant fully understands the described scenario and knows what to do. This means that we support them and answer potential questions if anything seems to be unclear. As mentioned in **C11** this can be a delicate matter. We have to take special care to provide enough information and make everything clear, but at the same time we must avoid to suggest any possible solutions or influence the results in any other way.

The results collected from the Drawing Study consist of the hand-drawn sketches and our personal notes. The purpose of the notes is to capture the intent of our participants throughout the design process. This helps us to understand what the drawings are supposed to convey and how more complex elements such as interaction or animation are incorporated. As already stated in **C4** this makes follow-up interviews obsolete, since we are already capturing all necessary information during the creation of the sketches.

Scenarios

The following four scenarios are presented to each participant in the same order as they are presented here. Each of them features a specific task that should be supported by the resulting visualization. The scenarios are chosen to be representative for specific tasks that might benefit from the visualization of uncertainties.

Scenario 1 - Bus Scenario
The first assignment is to create a visualization that supports the user in gauging the probability that an event will happen before/after a given point in time. The concrete scenario described to the participants is as follows: *'A bus should arrive at 12:00, but may be running late for up to 10 minutes. How would you visualize this scenario, so that you can estimate the probability of still catching the bus if you arrive at the bus station at a given point in time?'*

Scenario 2 - Project Scenario (1/2)
The second scenario is about the comparison of two events with uncertain end times. The assignment is to create a representation that makes it possible to see which of the two events will end earlier on average. The concrete scenario if formulated like this: *'There are two possible approaches to a given project. The first approach will take 20 to 28 days, while the second one will take 23 to 26 days. How would you visualize the scenario, so you can effectively judge which of the two approaches will on average lead to an earlier completion of the project?'*

Scenario 3 - Project Scenario (2/2)
The third assignment works with the same scenario as the second one, but a different user task should be supported by the visualization. Instead of judging the average completion time, the user should be able to distinguish which of the two events has a higher probability of having ended before a given point in time. The concrete scenario is formulated like this: *'Consider the same two project approaches as before and an additional given point in time. How would you visualize the scenario, so you can effectively gauge which approach is more likely to have finished until the given point in time?'* Since this scenario and the previous one are identical apart from their tasks, they are collectively referred to as the *project scenario*.

Scenario 4 - Lecture Scenario
The fourth and last scenario is about judging the probability of two events overlapping in time (i.e. taking place at the same time). The concrete scenario is formulated like this: *'Two lectures are taking place after each other. the first lecture will end between 11:50 and 12:05, while the second lecture will start between 12:00 and 12:15. How would you visualize the scenario to be able to judge the probability of an overlap of the two lectures? Furthermore, it should be possible to accurately judge the interval in which an overlap can take place.'*

Participants

The target user group of this thesis is the general public. Therefore, our study group has to be a heterogeneous user group consisting of people of different ages, gender and educational background. As mentioned in **C12** part of the selected participants may be students recruited from university, but it is important to find people with a different background to complement them.

In total we recruited 32 participants for the Drawing Study. 20 of them are male and 12 of them are female. 24 participants were under the age of 30, while 8 were older. We did not collect any additional background information from our participants, but as mentioned in **C7** every participant is known to us, which makes it easy for us to acquire more information should it be needed for further analysis.

3.3 Results and Generated Hypotheses

We collected one sketch for each Scenario (1, 2 and 4) per participant. Exceptions to this are one participant who could not provide any sketches, as well as two participants who have not provided sketches for the first and second scenario respectively. Scenario 3 yielded only 2 additional sketches, because it only featured a different user task compared to Scenario 2. All but two participants argued that their sketch for Scenario 2 was also applicable for Scenario 3. This is the reason why both Scenarios (2 and 3) share a collective name and are also evaluated together.

In total 93 sketches and corresponding descriptions were collected. To identify popular approaches we decided to analyze the collected material using an open coding approach. This means that in the first step appropriate categories need to be defined, by which every sketch can be classified. In our case these categories were defined by cooperatively going through the collected material to look for distinctive features. This iterative process of reviewing sketches and adapting classification criteria led to the following categories:

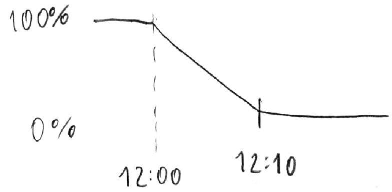

Figure 3.1: *This sketch shows an example of an explicit representation of uncertainty. The height of the graph directly encoded the probability of an event to the corresponding time.*

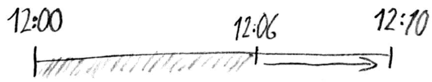

Figure 3.2: *This is a counterexample of the category of explicit uncertainty representations. The uncertain time frame between 12:00 and 12:10 is only marked by its bounds and no element in the visualization is directly mapped to the probability.*

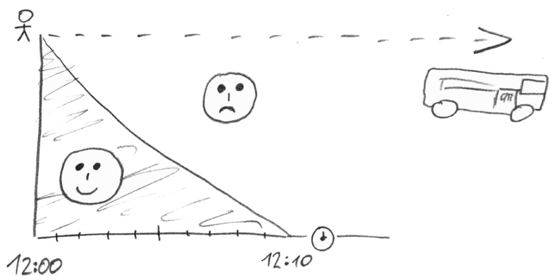

Figure 3.3: *In this sketch smiley faces are used to intuitively convey the notion of a 'good' or 'bad' outcome.*

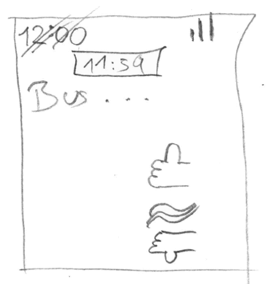

Figure 3.4: *This visualization design incorporates interaction and either provides the user with a thumbs-up or thumbs-down gesture conveying the probability of catching the bus.*

- **Cat. 1 – Explicit.** Sketches fall into this category if they feature some kind of explicit representation of uncertainty. This means that the uncertainty is not only given through the bounds of an interval, but also somehow directly encoded. Figure 3.1 shows an example of this category, while Figure 3.2 shows a counterexample. This category is further split up into four types of explicit uncertainty representations.
  - **Icons.** This category encompasses visualizations that encode uncertainty in some kind of icon. Examples of this are smiley faces as in Figure 3.3, or thumbs-up/down icons shown in Figure 3.4.
  - **Color Value.** Sketches that encode uncertainty in the color of an object fall into this category. Figure 3.5 shows an example.
  - **Length/Height.** If the uncertainty is conveyed through the length or height of an element, it falls into this category. An example can be seen in Figure 3.1.
  - **Interaction.** Some visualization designs incorporate user interaction to present uncertainty information for specific points in time more precisely. One of these designs is shown in Figure 3.4.
- **Cat. 2 – Temporal Line Chart.** If the visualization sketch features a conventional temporal line chart, it falls into this category. Figure 3.1 shows an example sketch.
- **Cat. 3 – Clock.** Drawings like the on shown in Figure 3.6 that feature a clock metaphor to convey the notion of time fall into this category.
- **Cat. 4 – Bounded.** Contrary to *Cat. 1* visualizations of this category convey uncertainty through the bounds of an interval. Figure 3.2 illustrates this with an example sketch.
- **Cat. 5 – Horizontal Time.** A sketch falls into this category if it features a horizontal time axis, as in Figure 3.5.
- **Cat. 6 – Vertical Time.** If the time is depicted on the vertical axis, the drawing falls into this category. An example of this can be seen in Figure 3.7.

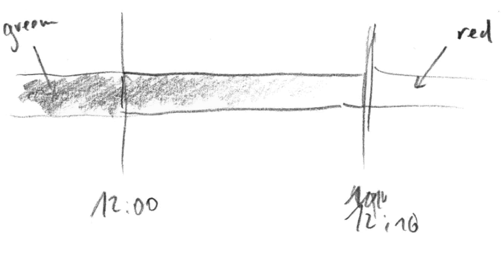

Figure 3.5: *This sketch is very similar to a Gradient Plot. The color fades from green to red to encode the probability of catching the bus.*

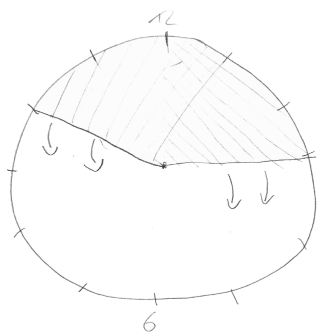

Figure 3.6: *This visualization utilizes a clock metaphor to convey the notion of time intuitively.*

Furthermore, the sketches of the *project scenario* and the *lecture scenario* are divided into **juxtaposed** and **superimposed** representations. **Juxtaposed** visualizations show the two events being compared in a side-by-side view, while **superimposed** approaches overlap both events in the same space. All sketches with superimposed views of the *lecture scenario* are further split up into a category that uses color to distinguish the two lectures and a category of drawings that does not.

Figure 3.7: *This sketch shows a rare example of a vertical time axis collected from our Drawing Study.*

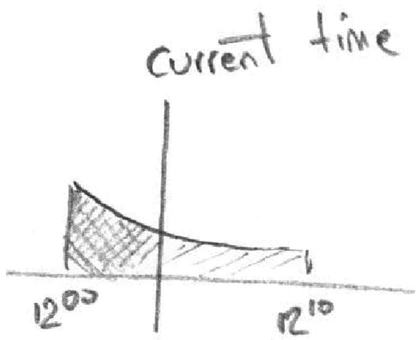

Figure 3.8: *This drawing illustrates that most categories are not mutually exclusive. It features a temporal line chart, as well as a color gradient to encode the uncertainty.*

It is important to note that not all of these categories are mutually exclusive. Therefore, a single sketch might count for multiple categories at once. Figure 3.8 illustrates this with an example.

The collected results of each scenario are presented in Figures 3.9, 3.10 and 3.11 respectively. The figures show bar charts with a bar for each category. The subcategories of the Explicit category and the division of the Superimposed category of the *lecture scenario* are represented through corresponding colors and textures.

As mentioned in **C6** the collected drawings are not analyzed in regard to the connection between the visualizations and the underlying data, as in Walny et al.'s [61] work. All further interpretation of the collected material stems from analyzing frequencies of different categories in the specific scenarios.

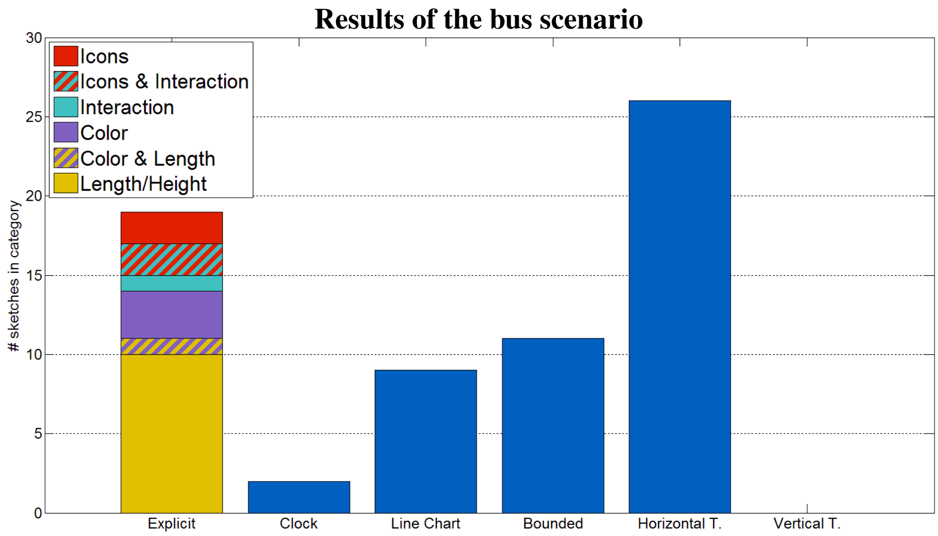

Figure 3.9: *These are the results of the bus scenario, which yielded 30 sketches in total. The bar chart shows how many sketches fall into each category. The category of explicit representations is further split into subcategories, which are distinguished by color. The hatched areas represent sketches that fall into multiple subcategories.*

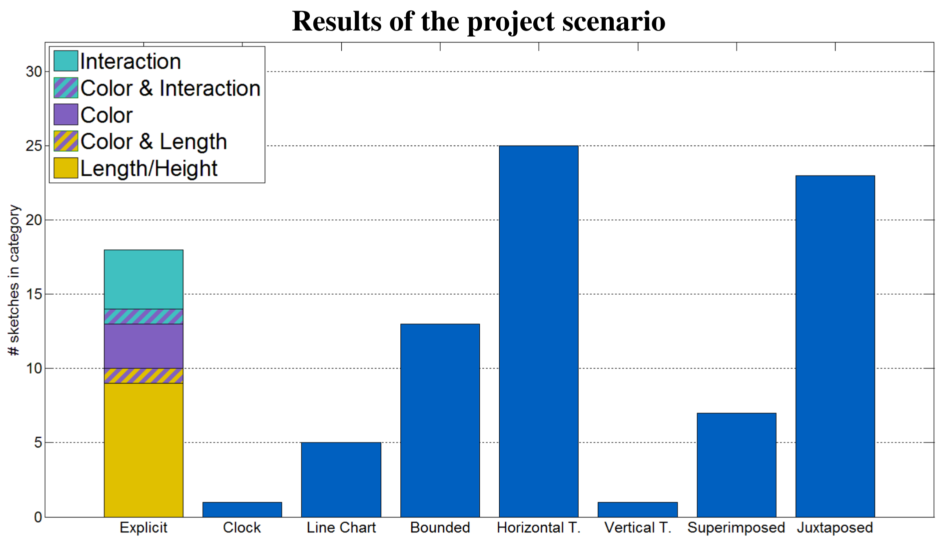

Figure 3.10: *These are the results of the project scenario, which yielded 32 drawings. The chart features the same categories as in Figure 3.9 and adds two additional ones to distinguish juxtaposed and superimposed representations.*

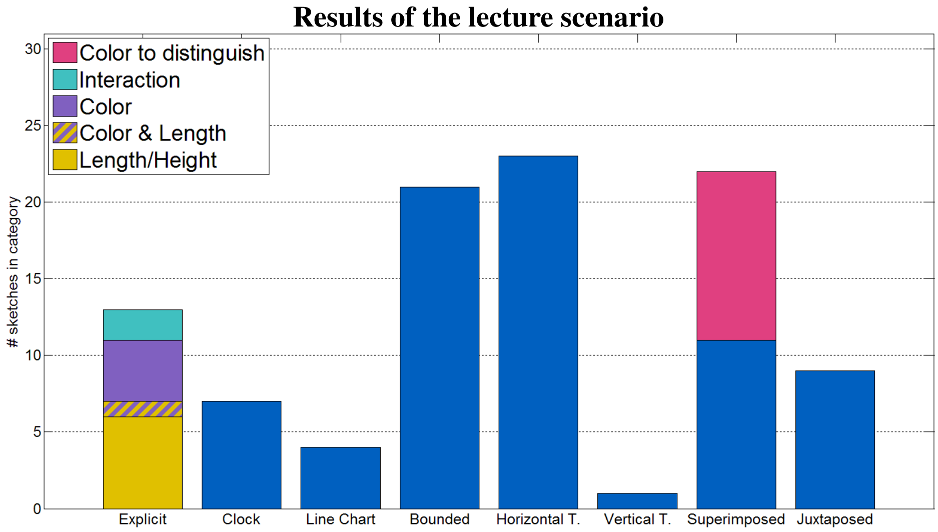

Figure 3.11: *The results of the lecture scenario are presented in the same way as the other scenarios. Furthermore, the category for superimposed representations is further split into those that use color to distinguish the visualized lectures and those that do not.*

Looking at the overall results it can be seen that uncertainty is most often represented through length or height. Icon representations, on the other hand, are only used in the *bus scenario*. This could be accounted for by it being the only scenario having a 'good' and a 'bad' outcome (catching or missing the bus), which lends itself to a visualization using smiley faces or a thumbs-up/down representation. Interaction is particularly popular in the *project scenario*. We believe that this is due to the comparison needed to solve the posed task. Through interaction the user can be provided with exact probability values, which can be compared accurately. In the *lecture scenario* the share of bounded representations is higher than in the other scenarios. We believe that this is due to the complexity of the task. It is hard to come up with a good visualization that actually conveys the overlap of two lectures well. This leads many participants to resort to a simple representation of two overlapping intervals. In the *project scenario* juxtaposed representations are more popular than superimposed views, while in the *lecture scenario* it is the other way around. A reason for this could be the nature and focus of the respective tasks. The *project scenario* focuses on the difference of the two intervals, while the *lecture scenario* specifically asks for the overlap of two intervals.

From the collected sketches, observations and category frequencies we derived 12 hypotheses which are presented in the following list:

- **DH1 - Temporal line plots are intuitive representations to support the user in judging a specific probability value of a given point in time.**
  The results of the *bus scenario* show that almost two thirds of all drawings feature an explicit representation of uncertainty of some kind. This makes sense, since the task description directly asked to support the user in determining the uncertainty at a given

point in time. In this context, especially temporal line plots, like the one in Figure 3.1, are common.

- **DH2 - Gradient Plots are intuitive representations to support the user in judging a specific probability value of a given point in time.**
  Another popular explicit uncertainty representation from our results is the Gradient Plot, like the one shown in Figure 3.5. This hypothesis is especially interesting in combination with the results of other studies, which is explained subsequently to this list.

- **DH3 - Icon representations, like smiley faces, are a good approach to represent probability values in a highly intuitive way, as long as these values do not have to be judged very precisely.**
  Other explicit representations of uncertainty utilize icons to convey probability values. An example of this can be seen in Figure 3.3. We believe this approach to be especially intuitive, but lacking the precision to represent exact values.

- **DH4 - Bounded visualizations are intuitive and effective ways to convey durations and temporal bounds of events with uncertain start and end times to non-expert users.**
  Even though most visualizations feature an explicit representation of uncertainty, bounded representations, like the example shown in Figure 3.2, are also often used. An important question that is left open by this observation is if bounded approaches were only used due to a lack of a better solution, or because they are really seen as a good approach. Either way, this approach seem to be intuitive to most people, even if it is not well suited for the task at hand. Gschwandtner et al.'s [22] results show that this approach is well suited to convey durations and temporal bounds to the user, which leads us to this joint hypothesis.

- **DH5 - It is more intuitive to a non-expert user group to vertically map time from the bottom to the top than vice versa.**
  The assumption that most people would represent time on the horizontal axis from left to right is supported by our results. From the total amount of 93 sketches we collected and analyzed, 80% represent time in this way. We also assumed that time would usually be represented from top to bottom if the time line occupies the vertical axis. This assumption does not seem to hold, since only two sketches feature time this way. One of them can be seen in Figure 3.12. The rest of the drawings show time either in a clockwise manner (clock metaphors) or from bottom to top, as in Figure 3.7. These unexpected findings could make the verification of this hypothesis interesting.

- **DH6 - If two or more events are compared to each other, it is more intuitive to show them in a juxtaposition than superimposed in the same space.**
  This hypothesis stems from the fact that there are more than three times as many juxtaposed approaches(23) than superimposed ones(7) for the *project scenario*.

- **DH7 - Icon representations are not well suited for direct comparison.**
  The results of the *project scenario* mostly feature explicit representations of uncertainty. Furthermore, there are multiple different approaches to convey uncertainty explicitly. Icon

representations are not used at all, though. We believe that this is due to the comparison task of this scenario which leads us to this hypothesis.

- **DH8 - Most people prefer to have the underlying uncertainty of data presented to them, even if it is not directly relevant for the task at hand.**
  The first task of the *project scenario* only called for a comparison of the mean values of the two uncertain events. The fact that many drawings still feature explicit representations of uncertainty leads us to this conclusion.

- **DH9 - To represent the amount of overlap between events, it is intuitive to superimpose them in the same view.**
  In contrast to the results of the *project scenario*, which featured more juxtaposed views than superimposed ones, the *lecture scenario* reversed this tendency. We believe this is due to the fact that the scenario described an overlap of two events and hence people drew the two events in an overlapping manner.

- **DH10 - Clocks lend themselves to show two superimposed time intervals, as long as the overlapping area does not exceed a one hour time frame.**
  The *lecture scenario* featured more clock metaphors than both other scenarios. This observation leads us to this hypothesis.

- **DH11 - Most people cannot think of an intuitive way to visualize the probability of two uncertain events to overlap each other.**
  The amount of explicit uncertainty representations is significantly lower in the *lecture scenario* than in both other scenarios. We believe this is due to the complexity of the task. Hence, most participants could not think of a good way to show the probability of an overlap in an explicit manner and resorted to a simpler, bounded visualization.

- **DH12 - Color is an intuitive way of separating two overlapping objects of the same shape.**
  Exactly half of all *lecture scenario* drawings that featured a superimposed view utilized color to distinguish the two overlapping events. An example of this can be seen in Figure 3.13.

It is important to note that not all of these hypotheses have been verified. They are merely meant to give direction to future research. Some of them build the basis for derived research questions for our User Survey explained in Chapter 4. The concrete research questions are presented in Section 4.2.

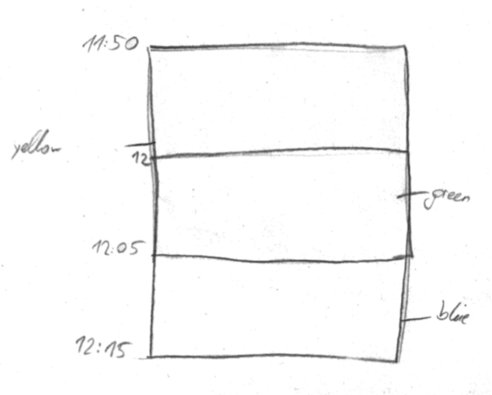

Figure 3.12: *This is one of only two collected drawings that features a vertical time axis which progresses from top to bottom.*

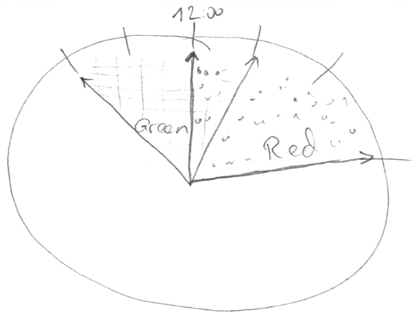

Figure 3.13: *This example shows a drawing that utilizes color to distinguish two overlapping intervals.*

Answers regarding the presented hypotheses are especially valuable in combination with the results of other works. Gschwandtner et al. [22], for instance, identified Gradient Plots to be well suited to judge specific probability values at given time points. However, it is not known how long it took the participants of their study to understand the visualization. Our results suggest that Gradient Plots are also intuitive and therefore readily comprehensible to most users. Hence, the technique seems to be a good choice to convey specific probability values to non-expert users. Moreover, a similar combined result can be drawn for ambiguation techniques. Gschwandtner et al. suggested to use them to convey bounds of temporal uncertainty and our results suggest that this approach is also intuitive. Corell and Gleicher [13] dedicated their work to highlight issues with error bar visualizations (see Figure 3.14). In this context we believe it is noteworthy that not a single collected sketch of our study contains anything resembling error bars.

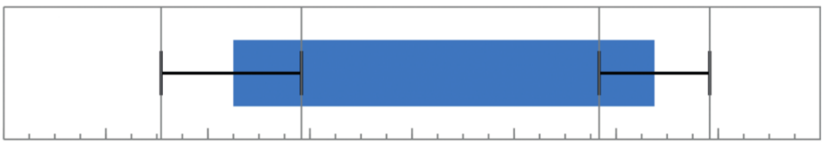

Figure 3.14: *This is a bar visualization utilizing error bars to convey temporal uncertainty. The blue bar reaches from the mean start time to the mean end time of the visualized event. The black error bars denote the bounds of the two uncertainty intervals. [22]*

CHAPTER 4

User Survey

This chapter presents the details of our User Survey in three sections. The introduction in the following section recaps the goal and our approach to achieving it.

In Section 4.2 the concrete research questions are presented. We illustrate how they are derived from the Drawing Study's results and discuss their relevance for the scientific community.

The last section is about the survey's design. The concrete approach and the reasoning behind it are thoroughly discussed. Furthermore, details about the implementation are given. Statistics about our participants are discussed at the end of this section.

The survey's results are presented in Chapter 5.

4.1 Introduction

To make the design and rationale behind the User Survey easily comprehensible, we recap on our goal and explain our approach. The previously presented Drawing Study constitutes the basis for this survey. It functions as a pre-study to identify interesting hypotheses to be tested. In total we listed 12 hypotheses, derived from the results of the Drawing Study (see DH1 - DH12 in the last chapter). In the next section the selection of hypotheses and the concrete research questions of this survey will be explained.

User studies with the goal of evaluating visualization approaches typically measure the accuracy and time taken by participants to solve a given task with the support of a visualization [36]. Since our goal is not to evaluate the efficiency and effectiveness of a technique, but how intuitive it is to the users, we adopt another approach of evaluation. To gauge the intuitiveness of different approaches, we directly ask our participants for their opinion on the matter. After all, our participants can give us direct feedback on whether or not the presented visualization makes sense to them and is understandable. In the categorization of Lam et al. [36] this approach falls under the category of *'Evaluating User Experience'*.

4.2 Research Questions

To define the research questions of this survey the results of the Drawing Study are reviewed and the most promising hypotheses are selected. Most promising in this context means that answers toward derived research questions lead to meaningful implications for the InfoVis community and for the design of future visualizations. Furthermore, the hypotheses are picked and derived into research questions in a way that leads to the comparison of two alternatives (e.g. visualizations with explicit uncertainty vs. visualizations without visible uncertainty). This helps in obtaining precise answers toward the research questions, which are well generalizable.

The following list presents the four final research questions of this survey derived from the Drawing Study's results. Furthermore, these questions also represent the main issue of this thesis:

- **RQ1 - Is it more intuitive to the average user to use Gradient Plots or temporal line charts to judge a specific probability value of an event at a given point in time?**
  The results of the Drawing Study suggest that Gradient Plots and temporal line plots are both intuitive approaches to convey the probability of an event at a given point in time (see **DH1** and **DH2**). We believe that it would be valuable to know which of the two approaches is generally perceived as more intuitive.

- **RQ2 - Is it more intuitive to the average user to visualize a comparison of two events, with uncertain temporal bounds, in a superimposed view (overlapping representation) or a juxtaposed view (side-by-side view)?**
  Judging from the Drawing Study, people seem to prefer juxtaposed visualizations to compare two events with uncertain start and/or end times (see **DH6**). Our subjective assessment of the collected sketches is that the superimposed approaches generally support the comparison task better, especially if the average durations of events are compared. Hence, we want to find out which of the two approaches is perceived as more intuitive.

- **RQ3 - Is it more intuitive to the average user to use an explicit uncertainty representation or uncertainty encoded in icons if the task at hand only calls for a rough approximation of the probability?**
  The Drawing Study suggests that icon representations are intuitive approaches to convey rough probabilities, but lack the precision to convey exact values (see **DH3**). This leads us to the question if people generally prefer to see a more exact representation of uncertainty or an icon representation if a rough probability estimate is sufficient for the task at hand.

- **RQ4 - Is it more intuitive to visualize an underlying uncertainty or to omit it if the uncertainty is not directly relevant for the task at hand?**
  The results of the Drawing Study suggest that most people like to see the underlying uncertainty of the duration of an event even though it is not relevant for the task at hand (see **DH8**). These results could be due to the fact that the study is aimed at the topic of uncertainty visualization. This could have induced people to visualize the uncertainty even though it is not needed for the task. Hence, we want to find out whether most people really want to see this underlying uncertainty or not.

A reason to select these three research questions is our believe that their answers will be valuable for future InfoVis design. As already mentioned in the results section of the Drawing Study chapter, results suggesting that Gradient Plots are intuitive could be very valuable, because of earlier results about this visualization technique. If they prove to be even more intuitive as simple temporal line plots, they would be very recommendable for future use. When designing a comparative visualization there is always a decision to be made between a juxtaposed or a superimposed approach. Answers to our research question **RQ2** could help future designers in this choice. Question **RQ3** has the potential to have a big impact on future visualization design. If our results suggest that most people are interested in being presented with the underlying uncertainty of visualized data, it would mean that many conventional techniques could be improved in this regard. Most deployed approaches do not deal with the notion of uncertainty if they do not specifically focus on this aspect of data [6]. Our results could suggest that it is important to think about visualizing uncertainty even if it is not the central focus of the visualization.

4.3 Design

As mentioned before, every research question is a comparison of two alternative visualization approaches for a given task. Hence, we compare the two approaches using a common scenario, which can be represented by both techniques. The scenarios are similarly structured as in the Drawing Study, providing context information, as well as a task to solve within the scenario.

Since we want to directly compare techniques to each other, but also want to get an overall rating of intuitiveness of every approach, we do not only ask for a binary decision between two alternatives. Our approach to elicit a numerical rating and also a clear decision for one of the two alternatives works as follows. First the scenario is presented with one of the two techniques. The participant is then asked to rate this first alternative on a Likert scale [38] from 1-10. This is followed up with a repetition of the same scenario and the same first visualization. Additionally, the second alternative is shown and the user is asked how it compares to the first one. The answer can be given on a Likert scale from 1-3 (Worse - Equal - Better). By randomizing which of the two visualizations is shown first and rated on the 1-10 scale and which comes second and is compared to the first one, we gather ratings and comparisons for every visualization over multiple participants.

The goal of these ratings is to determine how intuitive a visualization technique is to our participants in a specific scenario. To enable our participants to provide these ratings, it is important to make sure that they understand how the visualization works in principle. As mentioned in **C13**, we accomplish this through an introduction at the beginning of the survey. It presents all visualization techniques used (Gradient Plot, temporal line graph, juxtaposed/superimposed temporal line graph, icon visualization and bar chart without uncertainty) and gives example scenarios to illustrate their function. Furthermore, to make sure that participants do not mistakenly think they are understanding the visualization of a scenario and its solution, we provide the correct solution for each scenario at the end of the scenario description. This means that a participant who misunderstands a visualization and therefore comes up with a wrong solution to a scenario's task, the misunderstanding will be identified because the participant's solution does not match the correct solution given in the text.

Misunderstandings like this, other issues with the visualization or the scenarios, explanations of the given answer as well as any other remarks may be given textually through a provided text field after every question and after every part of the introduction. This form of feedback allows us to identify problems with the scenarios, tasks and visualizations and also gives us qualitative feedback to understand the ratings.

Figure 4.1 shows the structure of our User Survey. At first every participant is asked to provide some personal data (name, age, gender and level of education). As mentioned in **C7** we collect this data because it could be used to analyze the survey results further. In the next step the introduction of visualization techniques starts. There are four major techniques (Gradient Plot, temporal line chart, bar chart with icon representation and bar chart without uncertainty), which are presented one after the other. Every technique is textually and visually described and presented. Furthermore, there are examples to illustrate their function.

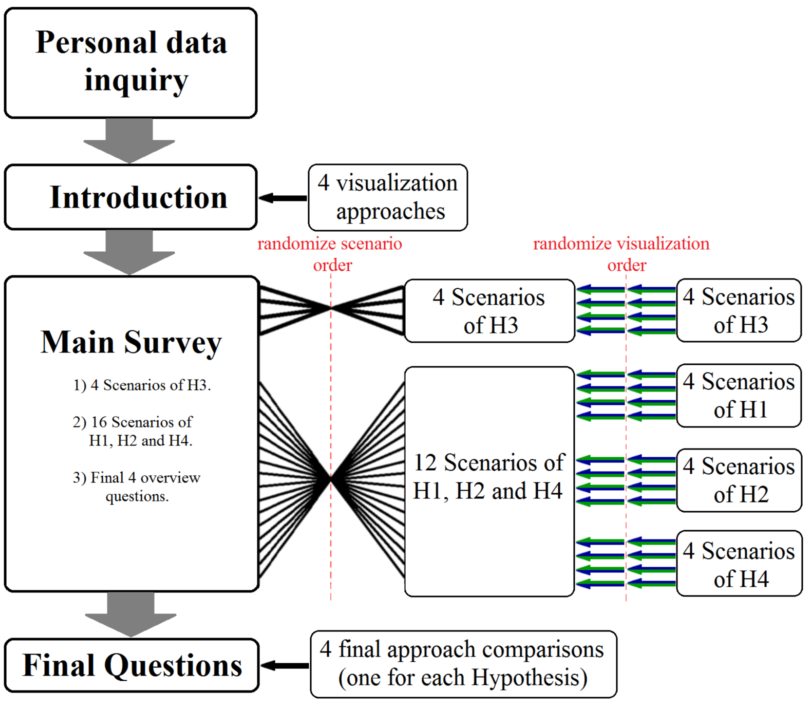

Figure 4.1: *This diagram illustrates the described structure of our User Survey. The left side shows the sequence of survey sections from the top to the bottom. The right elements illustrate which data is used in each section and how it is randomized in order.*

Following the introduction the four scenarios of **RQ3** are presented in a random order. The reason for presenting these scenarios before the other ones is that all tasks are about judging probabilities and sometimes comparing them. In this regard the tasks of the scenarios of **RQ3** are somewhat different, because they only ask for a rough estimation of a probability. For this reason we believe that it is better to not mix them in between other scenarios. The participants are also notified about these changing task characteristics. After these four scenarios the remaining 12 scenarios of the other research questions are presented. Again the sequence of scenarios is randomized per participant. Furthermore, the order of the two alternative visualizations per scenario is also random in every presented scenario.

After this main part of scenarios the participants are prompted with four final questions. These questions always appear in the same order. First the users are presented with a compilation of all visualizations of the scenarios of **RQ1**. On the left side all temporal line graphs are shown, while the right side shows all Gradient Plots (see Figure 4.2). The users are then asked for their opinion which of the two visualization approaches is generally more intuitive. The answer can be given through a Likert scale from 1-3 (left approach - equally intuitive - right approach). After this according compilations and questions are presented for the three remaining research questions. These four final comparisons conclude the survey.

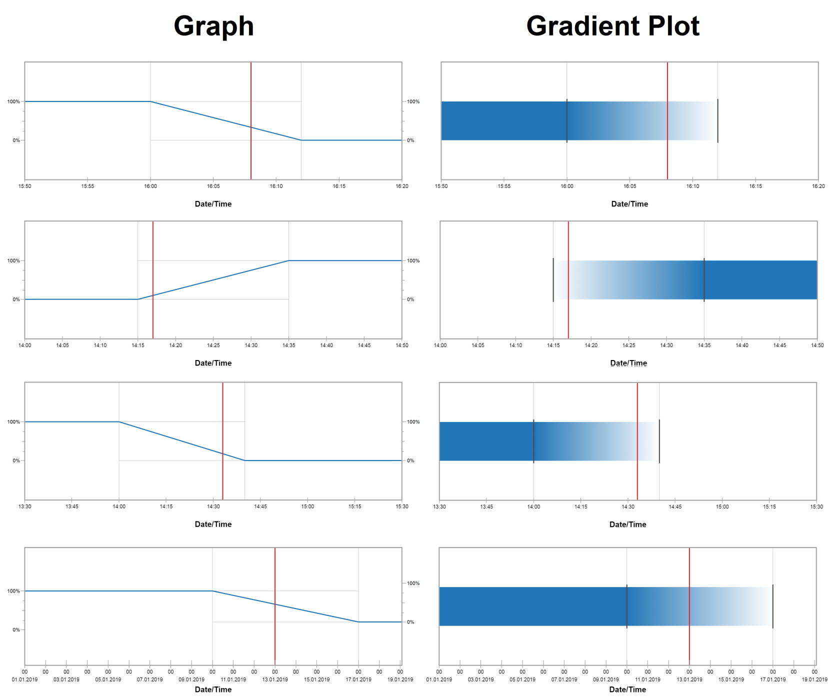

Figure 4.2: *Each line shows the two alternative visualizations for the four scenarios of* ***RQ1****. On the left are the temporal line charts and on the right are the Gradient Plots. In this final inquiry the participants are asked which of the two approaches is generally more intuitive to them.*

For our Drawing Study we recruited 32 participants. This is a sufficient number for a qualitative pre-study. The presented User Survey, however, aims to measure the intuitiveness of visualizations in a quantitative manner. This means that a greater number of participants are needed to collect enough data. For this reason we decided to take a different approach from the individual interview session of the Drawing Study. Meeting with every participant separately to conduct the User Survey simply would take too much time. Hence, the User Survey is implemented as an online questionnaire. This approach offers a series of additional benefits as well. By simply providing our participants with a link to the survey, they are free to complete it anytime. Furthermore, less organizational effort is required, since no meetings between us and our participants have to be scheduled. The online implementation also enables the questionnaire to be viewed and answered on many different devices, which further simplifies the conduct of the study. To implement the survey and make it easily accessible to our participants we utilized Google Forms [20]. To built the Google Forms with a randomized question order and randomized visualization order, as previously described, we utilized Google Apps Script [19].

Scenarios

For each of the four research questions there are four scenarios. This totals 16 scenarios for every participant. The following list describes the nature of the scenarios for each research question and presents an example scenario:

- **Scenarios of RQ1:** The scenarios and tasks of **RQ1** are very similar to the *bus scenario* of the Drawing Study. The participant is presented with an uncertain start or end interval of an event. The task is to determine the probability that the event has already started/ended at a given time point. The two visualizations provided for these scenarios are Gradient Plots and temporal line charts.
  This is one of the four scenarios as an example: 'A lecture is scheduled to start at 14:15, but may be delayed by up to 20 minutes. You want to estimate the probability of arriving too late if you will be there at 14:17. (10%)' Figures 4.3 and 4.4 show the two according visualizations.

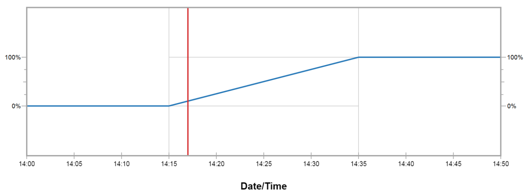

Figure 4.3: *This temporal line chart shows the probability of a lecture to have already started. Since uniform probability distributions are always assumed, the probability rises linearly from 0% to 100%.*

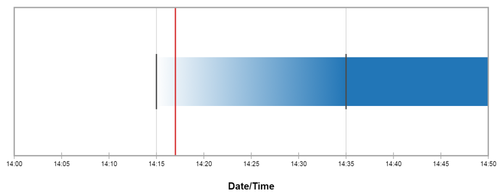

Figure 4.4: *This Gradient Plot shows the same lecture as Figure 4.3. The probability is encoded in the color value of the gradient, rising from 0% (white) to 100% (dark blue).*

- **Scenarios of RQ2:** To test whether people find juxtaposed or superimposed views more intuitive, we provide scenarios that compare two durations of events. The task is to determine which of the two events is more likely to have finished/not finished up to a given point in time. We visualize these scenarios with temporal line plots, which either show the two events juxtaposed or superimposed.
  This is an example scenario: 'You want to download a big file and can do that either via your Internet connection (A) at home, or via your mobile connection (B), which is supposed to be faster, but also more unreliable. The download will approximately take 30 to 40 minutes using your home connection and 24 to 42 minutes using your mobile connection. You need the file in 35 minutes and therefore want to pick the connection with the higher probability of finishing the download by then. (B)' Figures 4.5 and 4.6 show the two alternative visualizations for this scenario.

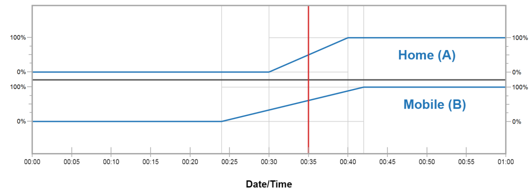

Figure 4.5: *This juxtaposition of temporal line charts shows a comparison of two Internet connections and the probability of when they will be done downloading a big file.*

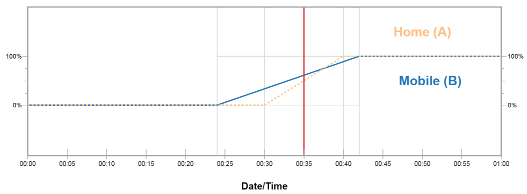

Figure 4.6: *This superposition of temporal line charts shows the same comparison as Figure 4.6. In contrast to the alternative visualization it overlaps both line charts in the same space.*

- **Scenarios of RQ3:** The scenarios for this research question are similar to the scenarios of **RQ1**. Their task is to estimate if an event will have started/ended before a given point in time. It is important to notice that contrary to the scenarios of **RQ1** a rough estimate of this probability is sufficient. The exact probability is not relevant. One of the two alternative visualizations is a Gradient Plot, while the other one is a simple bar visualizing the average start/end time. Additionally, the average visualization features a smiley face to encode the probability for the given point in time.
  This example illustrates the nature of these scenarios: 'You are sitting in a train, reading something on your phone. At 16:55 you will arrive at your destination. You want to roughly guess if your phone's battery will last until you arrive. (It is going to be very close, 33-65%)' The corresponding visualizations are presented in Figures 4.7 and 4.8.

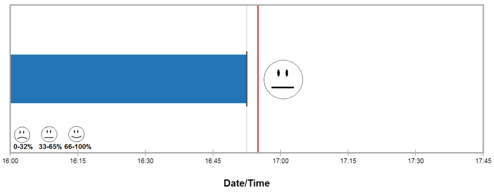

Figure 4.7: *This visualization shows the average value of the uncertain end time through a simple bar. The uncertainty for a given point in time (red line) is roughly represented by the smiley face, which can either smile, frown or look neutral.*

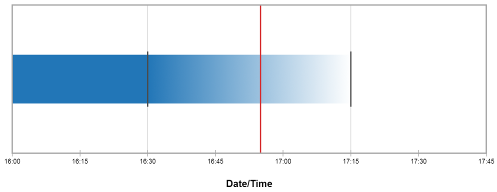

Figure 4.8: *This Gradient Plot shows the same scenario as Figure 4.7, but explicitly encodes the uncertainty over the whole uncertainty interval.*

- **Scenarios of RQ4:** These scenarios are derived from the first task of the *project scenario* of the Drawing Study. They compare two events with uncertain start/end times. The task is to determine which of the two events will start/end earlier/later on average. This means that the underlying uncertainty is actually not relevant to the task, as long as the average start/end time can easily be judged from the visualization. The compared visualizations are a bar showing the average start/end time only and a Gradient Plot with an added marker for the average start/end time.
  This is one of the concrete scenario descriptions: 'You want to download a movie you want to watch. There are two websites you can download it from, A and B. The download will take 17 to 24 minutes from website A, while it will take 18 to 29 minutes from website B. You want to determine which website to download from, to watch the movie as early as possible. (A)' Figures 4.9 and 4.10 show the two corresponding visualizations of the scenario.

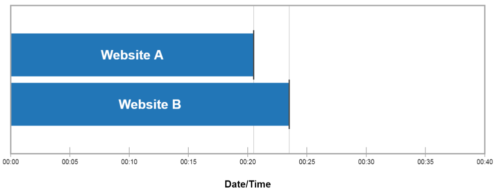

Figure 4.9: *This visualization only encodes the average values of the two durations. This information is enough to solve the task at hand.*

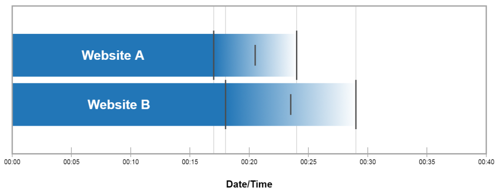

Figure 4.10: *This Gradient Plot features additional marks for the average values of durations, which are needed to solve the task of this scenario. The bounds of the uncertainty as well as the color gradient are additional information, which is not directly relevant for the task.*

4.4 Participants

The target user group of this survey is identical with the one of the Drawing Study - the general public. This means that we are again aiming for a heterogeneous group, consisting of people of various ages, genders, and educational background.

In total we privately recruited 60 participants from our personal social environment (i.e. family, friends, work colleagues, etc.). The gender distribution is exactly even, with 30 male and 30 female participants. With only 4 people not having completed Matura, 17 stating Matura as their highest education, 21 having completed a Bachelor's program as their highest education and 18 featuring a Master's degree or higher, our user group leans toward being well educated. The age distribution of the group looks like this: <20: 4, 20-24: 14, 25-29: 29, 30-34: 3, 35-39: 0, 40-49: 4, >50: 6.

CHAPTER 5

Results

In this chapter the results of our User Survey are discussed in detail. These results shed light on the four main research questions of this thesis. In the following section the research questions are separately recapped and all relevant data collected from the User Survey are presented.

In Section 5.5, the qualitative feedback, gathered in textual form during the survey, is discussed. This feedback offers a means to gain deeper insights about the reason for the quantitative results.

This chapter concludes with Section 5.6 which tackles the issue of limitations of our work. Furthermore, open questions which could be the focus of future work are discussed.

There are 4 scenarios and a final comparison regarding each research question. Each scenario represents a decision for the more intuitive one of two approaches. This means that from our 60 participants we collected a total of 240 decisions from scenarios and 60 additional decisions based on the overall comparison of the two approaches. For each research questions the frequencies of these decisions are presented. Furthermore, the survey yields numerical ratings from 1-10 for each visualization approach regarding a scenario type. Since the order of the two compared techniques is randomized per participant and scenario, we collected an unequal amount of ratings for each one. The number of ratings per visualization approach range from 108 to 132. These rating distributions are analyzed by a Welch's unequal variance t-test. This test allows us to judge if there is a significant statistical difference of mean values between two samples of normal distributions.

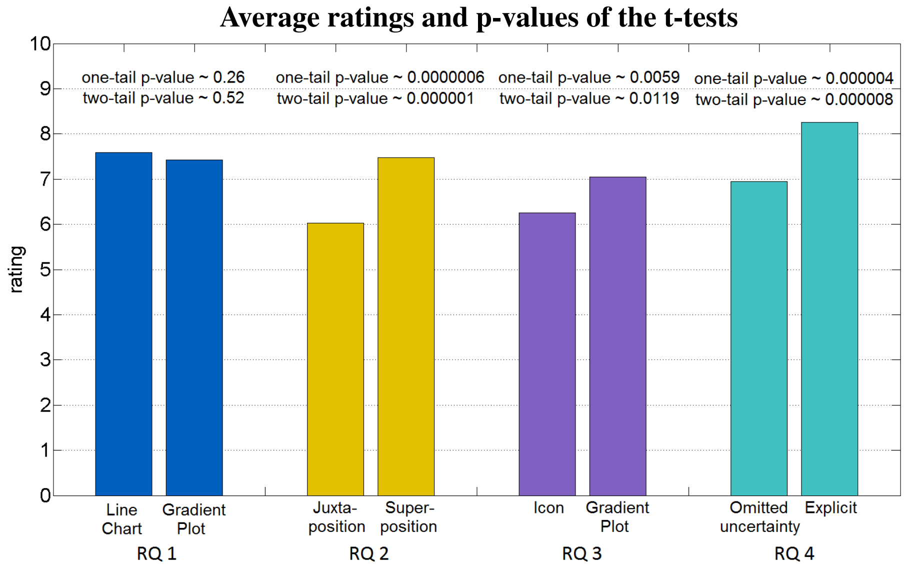

Figure 5.1: *This is a comparison of the average ratings of each visualization approach for a specific type of scenario. The first two bars show the visualizations of scenarios about RQ1. The next two are the ratings regarding RQ2 and so on. The p-values of the respective t-tests indicate the statistical significance of the rating difference (smaller is more significant).*

5.1 Research Question RQ1

RQ1 - Is it more intuitive to the average user to use Gradient Plots or temporal line charts to judge a specific probability value of an event at a given point in time?

The binary decisions for the more intuitive of the two approaches are presented in Figure 5.2. These results are not at all conclusive for either of the two approaches. The scenarios yielded 71/240 (30%) decisions for the temporal line chart to be more intuitive, 67/240 (28%) votes for the Gradient Plot and 102/240 (43%) of undecided votes. Furthermore, the results of the final comparison of the two techniques look similar. It features 20/60 (33%) votes for the line chart, 26/60 (43%) for the Gradient Plot and 14/60 (23%) of undecided votes.

A comparison of these results reveals that both techniques are similarly often perceived as more intuitive (30% - 28%, 33% - 43%). The main difference between scenarios and the final comparison is merely that more people decided for either one of the two approaches instead of giving an undecided vote. These results strongly indicate that neither of the two compared approaches is more intuitive than the other one.

This conclusion is further supported by the collected ratings and their t-test, which are presented in Figure 5.1. With respective average ratings of 7.59 (temporal line chart) and 7.43 (Gradient Plot), there is not much difference to be seen. Furthermore, the t-test yields results that do not allow us to reject the null-hypothesis (that both mean values are equal). This means that the two rating distributions cannot be distinguished with this sample size.

We therefore conclude with this answer toward our first research question: **Neither Gradient Plots nor temporal line charts are significantly more intuitive to judge a specific probability value of an event at a given point in time.**

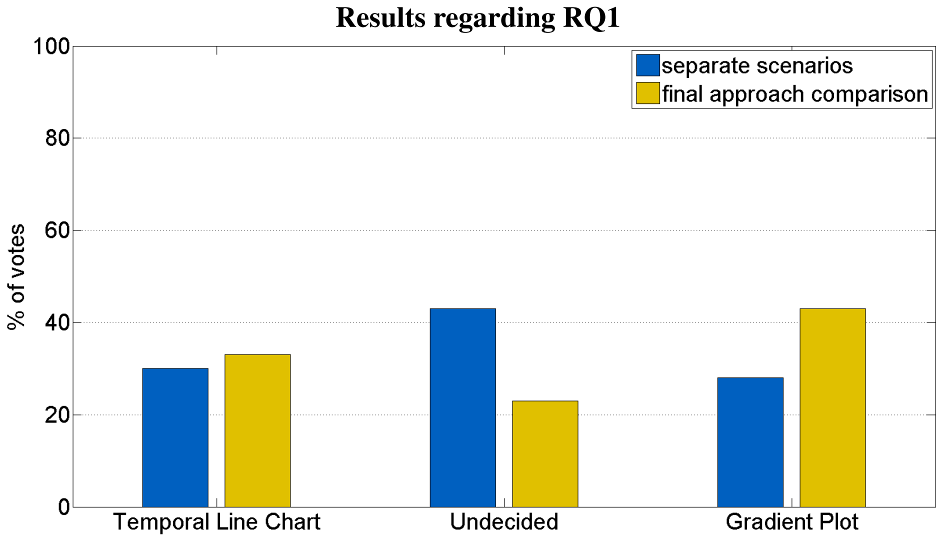

Figure 5.2: *These are the binary votes for when a decision has to be made which approach is more intuitive. The blue bars show the votes for each scenario, while the yellow bars indicate the votes in the final comparison at the end of the survey. This comparison is very even, with a similar amount of votes for each visualization technique and many undecided votes.*

Discussion of RQ1

Even though there is no clear indication which of the two techniques is the better choice, our results still offer valuable insights. By comparing the average scores attained by other techniques, it can be seen that both alternatives of RQ1 scored rather highly with scores of 7.59 and 7.43 respectively. This indicates that it should be possible to utilize either approach to intuitively and effectively represent temporal uncertainty to non-expert users.

Due to the indecisive results which of the two techniques should be chosen in a given scenario it probably comes down to the specifics of the area of application. One advantage of Gradient Plots is that they do not need any labeling of the y-axis, which makes the visualization more compact and arguably more esthetically pleasing. Conversely, some of our participants prefer temporal line plots because they feel they can perceive more precise probability values in this representation. Another argument could be made for temporal line plots, because most people learn how to read such visualizations in school and therefore have an easy time reading them, whereas Gradient Plots might be new to people and therefore harder to interpret at first.

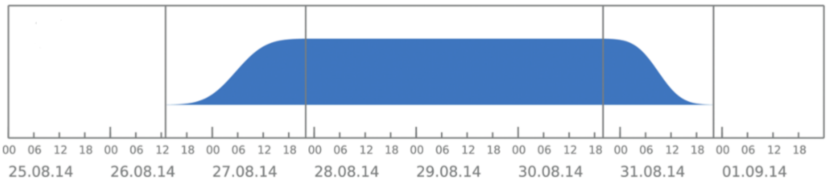

Figure 5.3: *The Accumulated probability visualization Gschwandtner et al. [22] test in their user study is very similar to the temporal line plots tested in our own work. It shows the probability accumulated over time of a certain event, which is identical to the data represented by our line plots. [22]*

Gschwandtner et al. [22] also compared Gradient Plots to a technique they call *Accumulated probability*, which is very similar to the temporal line plots we examined (see Figure 5.3). Based on their results Accumulated probability generally leads to a higher error rate when gauging specific probability values. Because of these findings and the argument that Gradient Plots can be represented more compactly, we recommend to use them over temporal line plots in most scenarios, unless the specifics of an application area favor temporal line plots.

5.2 Research Question RQ2

RQ2 - Is it more intuitive to the average user to visualize a comparison of two events, with uncertain temporal bounds, in a superimposed view (overlapping representation) or a juxtaposed view (side-by-side view)?

In contrast to the last one, the results about this research questions are very decisive toward one approach. As Figure 5.4 shows, there are only 28/240 (12%) votes for juxtaposition to be more intuitive from the separate scenarios. Superposition has 169/240 (12%) votes and 43/240 (18%) votes are undecided.

The final comparison question yielded very similar results. 12% (7/60) of all participants see juxtaposition as more intuitive, while 82% (49/60) prefer a superimposed view. 4% (7/60) participants were undecided in the final comparison.

The presented result clearly suggests that a superimposed representation is more intuitive. This is also supported by the collected ratings. Figure 5.1 shows that the average ratings of 6.03 (juxtaposition) and 7.48 feature a significant difference. The calculated p-values of the t-test are 0.0000006 and 0.000001, which shows that this result is also statistically significant.

This leads us to this conclusion toward our second research question: **For the average user it is more intuitive to visualize two events with uncertain temporal bounds in a superimposed view than in a juxtaposed view.**

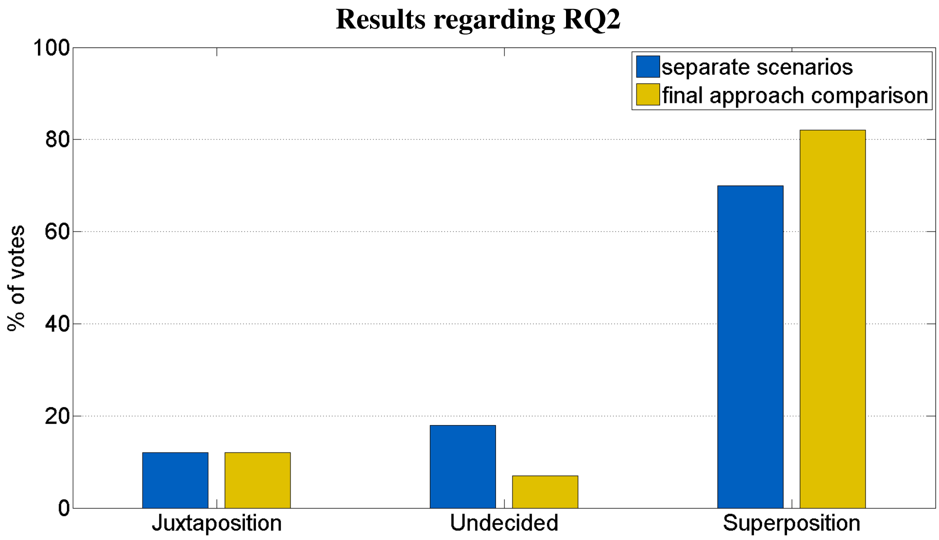

Figure 5.4: *The distribution of votes clearly indicates that superimposed representations are generally seen as more intuitive. During the scenario, as well as during the final comparison superposition was significantly more often seen as more intuitive. Only a comparatively small number of participants were undecided or saw juxtaposition as the superior approach.*

Discussion of RQ2

The qualitative feedback collected suggests minor shortcomings in the design of our juxtaposed representation technique. A major problem is the fact that the two line plots are not centered within their respective rectangle, but always shifted toward the middle line, separating the two plots. Furthermore, the visualization features many helper lines, which might clutter the representation too much. These flaws probably impact the final results to a small degree and lead to decreased ratings of this technique. Nevertheless, we believe that our superimposed view would still outperform an improved version of the juxtaposed one.

The results our study yielded also make a lot of sense. Since the corresponding scenarios all aim at the comparison of two values at the same position in time, it is no surprise that superimposed views fare better. This direct comparison becomes very clear and easy if the values are represented in the same space, as opposed to a juxtaposed view. As described in Chapter 3 about our pre-study, the idea to compare these two techniques stems from the fact that many participants generated sketches depicting juxtaposed views for similar scenarios. It is interesting to see that juxtaposition is often utilized by non-experts in their own visualization designs, but still seen as inferior to superposition in a direct comparison.

5.3 Research Question RQ3

RQ3 - Is it more intuitive to the average user to use an explicit uncertainty representation or uncertainty encoded in icons if the task at hand only calls for a rough approximation of

the probability?

The decision between an icon representation and an explicit representation is also relatively clear. Only 21% (51/240) of all scenarios are voted to be more intuitively visualized by the icon representation. In 57% (136/240) of all scenarios a Gradient Plot was seen as a superior solution. 22% of all decisions from the scenarios are undecided.

As in previously presented results these percentages are similar between the separate scenarios and the final comparison. The final comparison yielded 11/60 (18%) votes for the icon representation, 44/60 (73%) votes for the Gradient Plot and 5/60 (8%) undecided votes.

The consequent conclusion that explicit uncertainty representations are more intuitive than approaches using icons is also supported by the t-test using the collected ratings (see Figure 5.1). The icon representations are averagely rated with 6.25 points, while the compared Gradient Plots reached 7.05 points on average. The test's resulting p-values of 0.0059 and 0.0119 suggest that this difference is statistically significant.

Hence, we conclude with the following answer toward our research question: **Even if only a rough approximation of the probability is sufficient for a given task, an explicit encoding of uncertainty is more intuitive than uncertainty encoded in icons.**

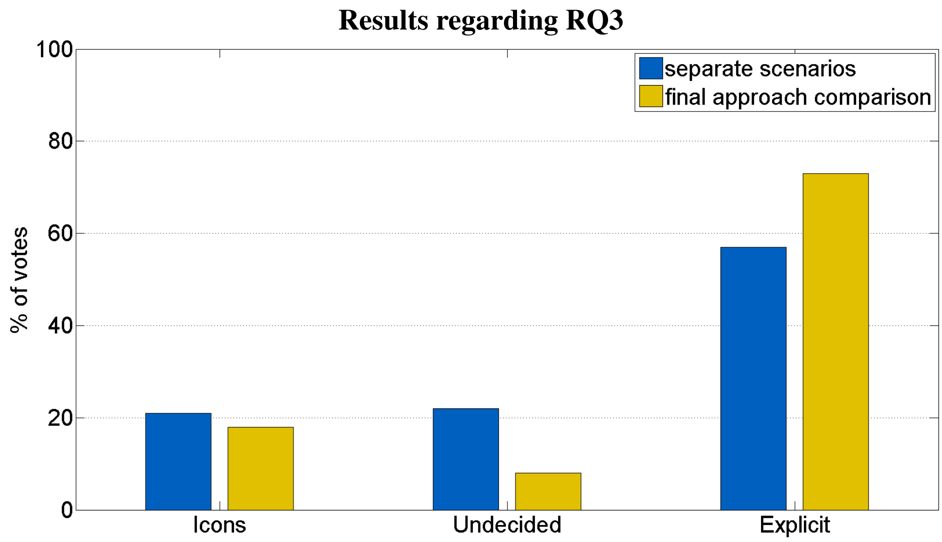

Figure 5.5: *It is clearly visible that most participants see an explicit representation of uncertainty as more intuitive than icon representations. From the scenarios this tendency is already obvious and is even more pronounced in the final comparison of the two techniques.*

Discussion of RQ3

The qualitative feedback shows that the design of our icon representations has some shortcomings. The main issue is that 0 percent probabilities share the same icon with low probabilities and high probabilities share the same with 100 percent probabilities. I. e. there are no separate representations for certain outcomes. Our design also only features three different icons, since they are only supposed to give a rough estimation of a value. It would be interesting to repeat the experiment with a new design that directly and gradually maps the probability values to a smiley face. This way both approaches (icons vs. explicit representation) would visualize the same information without aggregating or simplifying it. Some participants state in their qualitative feedback that icons are a good way to get a quick estimation, but they ultimately prefer an explicit approach because it is more precise. A continuously mapped smiley could possibly unite both of these qualities.

The benefit of icon representations to quickly and intuitively give the user a good estimate could be used to enhance more precise (but slower) visualizations. One could, for instance, use an icon additionally to a Gradient Plot. This way a user can get a quick estimation via the icon and more exact information from the Gradient Plot.

We believe that a major issue with icon representations is that they are highly dependent on the scenario of application. In the case of our study each scenario features a desired and an unwanted outcome. Therefore, smiling and frowning faces are a good choice. The same icons might not be fitting for different scenarios. Even if a scenario features a desired outcome, it may vary if a high probability or a low probability is to be fancied. This means that an icon representations always needs to be fitted to the scenario to be applicable. We believe that there is no good generic icon design that fits well for many cases.

5.4 Research Question RQ4

RQ4 - Is it more intuitive to visualize an underlying uncertainty or to omit it if the uncertainty is not directly relevant fo the task at hand?

The results regarding the fourth research question look similar to the ones of the previous two. The scenarios yielded 56/240 (23%) decisions for the visualizations without uncertainty to be more intuitive and 131/240 (55%) decisions for the explicit representation. 53/240 (22%) scenarios are undecided between the two approaches.

The results of the final comparison at the end of the survey again strongly correlate with the results of the scenarios. 14/60 (23%) participants see a visualization without uncertainty as more intuitive, while 37/60 (62%) believe the explicit approach is superior. 9/60 (15%) participants are undecided.

The explicit representation is also favored in the numerical ratings. While omitting the uncertainty yielded an average rating of 6.95, the Gradient Plots with additional average marker are rated 8.26 on average. This result is statistically significant with p-values of 0.000004 and 0.000008 from the t-test.

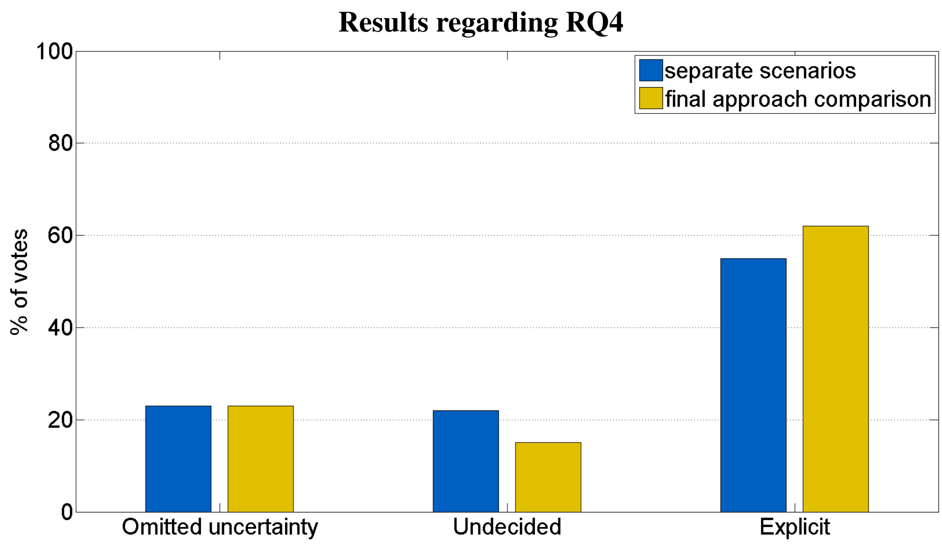

Figure 5.6: *The results of the fourth research question show a clear tendency toward visualizations with explicitly represented uncertainty. More than twice as many participants voted for this option in contrast to the visualizations omitting uncertainty.*

Discussion of RQ4

To us these results are the most interesting ones, as they have the greatest potential to have an impact on future visualization design. Even though all scenarios of our study are designed to be solvable without taking the underlying uncertainty into account, an overwhelming majority of participants prefers to have uncertainty visualized. These findings are so significant because most real world visualizations do not convey uncertainty, even though almost all represented datasets are afflicted by some kind of uncertainty (see 1.1 for further detail). Hence, the intuitiveness of countless visualization designs could possibly be enhanced by incorporating uncertainty into the visual representation.

It is important to note that this does not mean every visualization should deal with uncertainty for intuitiveness' sake. This additional data dimension usually takes up screen space and adds more visual complexity. This means that designers have to carefully consider the benefit of incorporating uncertainty into their design. Based on our findings, though, this consideration is not only important when uncertainty plays a major role in the scenario at hand. As our experiment shows there might be circumstances in which showing uncertainty, even though it is not vital for the task at hand, is advisable to provide the user with more information. This additional information may convey the feeling of being able to make better informed decisions and therefore lead to higher user confidence and satisfaction.

5.5 Qualitative user feedback

Qualitative feedback was collected textually during the survey. Corresponding text fields are provided at the end of each of the four visualization introductions. Through them participants can communicate any issues with their understanding of the presented techniques and/or give further remarks regarding them. Furthermore, after every question during the scenarios we also provide a text field for feedback. Our approach to analyzing the collected feedback is to simply find remarks that occurred multiple times and seem relevant to the quantitative results.

The most common feedback regards the intuitiveness of the icon representation. In total, eight participants state that they have some issue understanding the approach or explicitly state that they do not like the visualization because of a specific shortcoming. Three participants reason that the main problem is that there are no different icons for a 0% probability and a low probability up to 33%. The same is true for high probabilities and 100% probabilities. Also regarding the icon visualization, two participants state that they think that icons are very quick to read, but still prefer an explicit approach, because it is more precise and gives more information.

The second most common remark regards the orientation of the probability distributions, which was mentioned by five participants. If a visualization shows the extent of a car drive, for instance, it is more intuitive for some persons to show the probability of the drive still lasting (i.e. the probability decreases around the estimated arrival time). For other people it might be more intuitive to encode the probability of having already arrived (i.e. the probability increases around the estimated arrival time). We consistently encoded the probabilities of events to still be going on, which is not always the most intuitive solution for all participants, hence some remarked that a flipped visualization would be more convenient for them.

The superimposed temporal line charts work well for the comparison of two events, but might not be a good solution if there are more than two events to be shown. This remark is only stated once, but we believe that this is an important input. While superimposed charts are highly rated in our results, it is still important to keep in mind that such a visualization can quickly become cluttered if there is too much information overlapping in the same space.

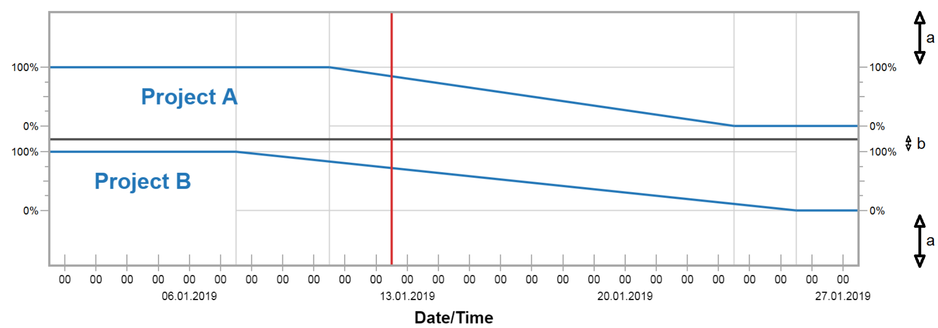

Figure 5.7: *On the right side of the figure the distances a and b are marked. These are the distances of the graph to its respective borders. The inequality of a and b confused some participants, which probably skewed the results toward the alternative visualization.*

Another feedback regarding the same research question of juxtaposition versus superposition, is critique on the juxtaposed visualization. The way we implemented this view features a different vertical distance from the bottom of one axis to the bottom of the frame and the top of the same axis to the top of the frame. I. e. the axes are not vertically centered within their sub-frame (see Figure 5.7). This confused at least four of our participants, who state this in their feedback. Shortcomings like this could skew our results to a certain degree, but we believe that even better implementations of all utilized techniques would have led to the same outcome.

Our results show that there is no clear favorite among temporal line charts and Gradient Plots, which is also reflected in the qualitative feedback. Two participants state that they prefer line charts because they yield more precise probability values. Four other participants argue that Gradient Plots are superior because they are more convenient to use. This is due to the obsolete y-axis. The desired probability value can be directly estimated from the color gradient.

In the comparison of RQ4, which features an approach without uncertainty and one with explicitly encoded uncertainty, our results show a clear tendency toward the explicit approach. The qualitative feedback of two participants gives an indication why this is the case. They argue that the additional information can be important in many real world scenarios and does not distract from the task of comparing mean values.

5.6 Limitations and future work

Obviously user studies like ours are always subject to certain limitations, which cap the generalization of results. Experiments are only samples taken of specific configurations, under specific conditions. To gain further insights and to make results applicable more generally, further evaluations need to be conducted. This future work can alter details of existing work and put its focus on other elements of the experiment to facilitate knowledge gain in another direction.

To keep the scenarios of our study as simple as possible to avoid distracting our participants by unnecessary details, we only worked with uniform probability distributions. In real world scenarios, though, countless other distributions might occur. Changing this underlying data property might also have an impact on the relative effectiveness and intuitiveness of the tested visualization approaches. This could be an interesting basis for future research.

Another restriction made for simplicity's sake is the lack of interaction. All visualizations of our User Survey are strictly static and cannot be interacted with. It would be interesting, however, to find out which kind of interaction works well in the context of conveying temporal uncertainty. Particularly how, how interaction can be offered in an intuitive way to novice users could be the focus of future work.

As already mentioned in the discussion of RQ3, the effectiveness of icon representations is dependent on the applicability of the used icons. Furthermore, other shortcomings of our icon design are discussed. Future work could focus on fixing these mistakes to conduct a similar study about icon visualizations for temporal uncertainty. We believe that well designed representations could indeed be very intuitive to non-expert users. Especially icons in addition to an approach depicting an explicit encoding of probability values could combine the benefits of both.

Our Drawing Study yields a list of 12 hypotheses. Five of them are tackled by our User Survey. The remaining ones are still open and unverified statements. Evaluating them in further user studies could yield valuable insights into the visualization of temporal uncertainty.

CHAPTER 6

Conclusion

Most data collected from the real world contains some kind of uncertainty. There are many reasons for this which lead to different types of uncertainty. An overview of these types can be found in Gschwandtner et al. [22]'s work. In recent years there has been an effort to incorporate this data property into visual representations. This provides the user with more information and a better understanding of the data, which in turn enables the user to make better informed decisions. All this is also true for the field of InfoVis of time-oriented data.

Especially for non-expert users it is important to provide intuitive visualizations for them to be able to effectively utilize it. In regard to temporal uncertainty there is still a lot of work to be done to find out which visualizations are the most intuitive ones. To gain deeper insights into this matter, user studies are generally a good approach. Since intuitiveness is such a subjective issue, user studies provide the means of evaluating subjective opinions of participants and objectify them through a large enough sample size. A downside of user studies is that they can generally answer only rather specific questions [33]. Hence, it is important to first identify interesting research questions before evaluating them in a comprehensive study. This is the reason we conducted two user studies. The first study, called the Drawing Study, acts as an exploratory pre-study, which yields possible research questions. The second study, called the User Survey, then evaluates research questions and yields the desired insights.

To provide a good overview of the state-of-the-art of temporal uncertainty visualization a comprehensive literature research was conducted. We present a multitude of example projects which focus on representing temporal uncertainty. Furthermore, to provide the necessary theoretical background for the practical part of our work, we present the state-of-the-art of user studies in the field of InfoVis. This current state is represented by a selection of presented example projects and additionally by works that focus on the theory behind user studies.

The first study we conducted is called the Drawing Study because we asked our participants to draw their own visualization sketches. These sketches are based on provided scenarios about temporal uncertainty. Overall, there are four scenarios, featuring different user tasks that should be supported by the visualization. The collected drawings are analyzed through an open coding approach. This means that they are categorized by similarities in their approaches and the total number of drawings falling into each category is interpreted. Based on this analysis we generated a list of 12 hypotheses regarding the intuitiveness of temporal uncertainty visualizations. Some of these hypotheses are not further addressed in this work and could be the basis for future research.

The most promising hypotheses of our list were formulated into concrete research questions, which we evaluated in our main study, the User Survey. This study is implemented as an online survey utilizing Google Forms [20]. Each of the 60 participants completed the survey separately. Each participant was presented with 16 scenarios, four for each of the 4 research questions, in total. Each scenario featured a scenario text, describing a situation, and a specific user task that should be fulfilled. At first each participant was presented with one of two visualizations and asked to rate how intuitively it supports the task. Then the other visualization was presented as an alternative and the participant was asked to rate it as better, equally good or worse than the first one. The User Survey concludes with four final comparisons, one for each research questions. Each comparison features all corresponding visualizations presented in the scenarios before and the participant is asked to make a final decision on which visualization technique is more intuitive overall.

The results of the User Survey were analyzed using a Welch's unequal variance t-test to find statistically significant differences between ratings of compared visualization techniques. Furthermore, the binary decisions (i.e. which visualization approach is more intuitive) gathered from each scenario additionally show the general tendency of our participants. The results and their statistical analysis are thoroughly discussed in Chapter 5. One of the most intriguing results is that icon visualizations performed relatively poorly in our survey. This comes as a surprise to us, because icons seem to be very intuitive at first glance. A reason why they did not perform better could be that icons need to be specifically designed for their situated use. An icon representation is most intuitively understandable if it fits the scenario and task at hand very well. In our case we worked with generic smiley faces, which could have been too general to work well. These open questions about icon visualizations offer opportunities for future research. Another interesting finding is that our participants preferred to have uncertainty visualized to them, even if it was not relevant for the task at hand. This is an important finding, because it could impact future visualization design. Most InfoVis systems are not designed to incorporate uncertainty into their data representations, but our results indicate that users generally like to have this information presented to them additionally. Hence, existing visualization systems potentially could be improved by incorporating uncertainty information and future designers should think about uncertainty even if it is not central to the task at hand.

Bibliography

[1] Aerts, J. C., Clarke, K. C., and Keuper, A. D. (2003). Testing popular visualization techniques for representing model uncertainty. *Cartography and Geographic Information Science*, 30(3):249–261.

[2] Aigner, W., Hoffmann, S., and Rind, A. (2013). EvalBench: A Software Library for Visualization Evaluation. *Computer Graphics Forum*, 32(3pt1):41–50.

[3] Aigner, W., Miksch, S., Müller, W., Schumann, H., and Tominski, C. (2007). Visualizing time-oriented data—a systematic view. *Computers & Graphics*, 31(3):401–409.

[4] Aigner, W., Miksch, S., Thurnher, B., and Biffl, S. (2005). PlanningLines: Novel Glyphs for Representing Temporal Uncertainties and their Evaluation. In *Proceedings of the 9th International Conference on Information Visualisation (IV 2005)*, page 457–463. IEEE Computer Society Press, IEEE Computer Society Press.

[5] Alberti, K. (2013). Web-based visualization of uncertain spatiotemporal data. Master's thesis. Utrecht University.

[6] Amar, R. and Stasko, J. (2004). Best paper: A knowledge task-based framework for design and evaluation of information visualizations. In *Proceedings of the IEEE Symposium on Information Visualization*, INFOVIS '04, pages 143–150. IEEE Computer Society.

[7] Association for Computing Machinery (2018). Acm digital library. `https://dl.acm.org/`. Accessed: 2018-03-14.

[8] Benjamini, Y. (1988). Opening the box of a boxplot. *The American Statistician*, 42(4):257–262.

[9] Brodlie, K., Osorio, R. A., and Lopes, A. (2012). A review of uncertainty in data visualization. In *Expanding the frontiers of visual analytics and visualization*, pages 81–109. Springer Verlag London.

[10] Chittaro, L. and Combi, C. (2001). Visual definition of temporal clinical abstractions: A user interface based on novel metaphors. In *Conference on Artificial Intelligence in Medicine in Europe*, volume 2101, pages 227–230. Springer.

[11] Chittaro, L. and Combi, C. (2003). Visualizing queries on databases of temporal histories: new metaphors and their evaluation. *Data & Knowledge Engineering*, 44(2):239–264.

[12] Choonpradub, C. and McNeil, D. (2005). Can the box plot be improved. *Songklanakarin Journal of Science and Technology*, 27(3):649–657.

[13] Correll, M. and Gleicher, M. (2014). Error bars considered harmful: Exploring alternate encodings for mean and error. *IEEE transactions on visualization and computer graphics*, 20(12):2142–2151.

[14] Drecki, I. (2002). Visualization of uncertainty in geographical data. In *Spatial Data Quality*, pages 140–159. Taylor & Francis, London.

[15] Edwards, L. D. and Nelson, E. S. (2001). Visualizing data certainty: A case study using graduated circle maps. *Cartographic Perspectives*, (38):19–36.

[16] Ferris, B., Watkins, K., and Borning, A. (2010). Onebusaway: results from providing real-time arrival information for public transit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1807–1816. ACM.

[17] Gerharz, L., Pebesma, E., and Hecking, H. (2010). Visualizing uncertainty in spatio-temporal data. In *Spatial Accuracy 2010*, pages 169–172.

[18] Gomez, S. R., Jianu, R., Ziemkiewicz, C., Guo, H., and Laidlaw, D. (2012). Different strokes for different folks: visual presentation design between disciplines. *IEEE transactions on visualization and computer graphics*, 18(12):2411–2420.

[19] Google LLC (2018a). Google apps script. `https://developers.google.com/apps-script/`. Accessed: 2018-03-14.

[20] Google LLC (2018b). Google forms. `https://www.google.com/forms/about/`. Accessed: 2018-03-14.

[21] Google LLC (2018c). Google scholar. `https://scholar.google.com/`. Accessed: 2018-03-14.

[22] Gschwandtner, T., Bögl, M., Federico, P., and Miksch, S. (2016). Visual encodings of temporal uncertainty: A comparative user study. *IEEE transactions on visualization and computer graphics*, 22(1):539–548.

[23] Haemer, K. W. (1948). Range-bar charts. *The American Statistician*, 2(2):23–23.

[24] Harris, R. L. (2000). *Information graphics: A comprehensive illustrated reference*. Oxford University Press.

[25] Heer, J., Kong, N., and Agrawala, M. (2009). Sizing the horizon: the effects of chart size and layering on the graphical perception of time series visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1303–1312. ACM.

[26] Hintze, J. L. and Nelson, R. D. (1998). Violin plots: a box plot-density trace synergism. *The American Statistician*, 52(2):181–184.

[27] Hullman, J. (2016). Why evaluating uncertainty visualization is error prone. In *Proceedings of the Beyond Time and Errors on Novel Evaluation Methods for Visualization*, pages 143–151. ACM.

[28] IEEE (2018). Ieee xplore digital library. `https://ieeexplore.ieee.org/Xplore/home.jsp`. Accessed: 2018-03-14.

[29] Kao, D., Luo, A., Dungan, J. L., and Pang, A. (2002). Visualizing spatially varying distribution data. In *Proceedings of the Sixth International Conference on Information Visualisation, 2002.*, pages 219–225. IEEE.

[30] Kardos, J., Moore, A., and Benwell, G. L. (2003). The visualisation of uncertainty in spatially-referenced attribute data using trustworthy data structures. *15th Annual Colloquium of the Spatial Information Research Centre*, 15:11–26.

[31] Kay, M., Kola, T., Hullman, J. R., and Munson, S. A. (2016). When (ish) is my bus?: User-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5092–5103. ACM.

[32] Kinkeldey, C., MacEachren, A. M., and Schiewe, J. (2014). How to assess visual communication of uncertainty? a systematic review of geospatial uncertainty visualisation user studies. *The Cartographic Journal*, 51(4):372–386.

[33] Kosara, R., Healey, C. G., Interrante, V., Laidlaw, D. H., and Ware, C. (2003). Thoughts on user studies: Why, how, and when. *IEEE Computer Graphics and Applications*, 23(4):20–25.

[34] Kosara, R. and Miksch, S. (2001). Metaphors of movement: a visualization and user interface for time-oriented, skeletal plans. *Artificial intelligence in medicine*, 22(2):111–131.

[35] Kunz, M., Grêt-Regamey, A., and Hurni, L. (2011). Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system. *Natural hazards*, 59(3):1735–1751.

[36] Lam, H., Bertini, E., Isenberg, P., Plaisant, C., and Carpendale, S. (2012). Empirical studies in information visualization: Seven scenarios. *IEEE transactions on visualization and computer graphics*, 18(9):1520–1536.

[37] Leitner, M. and Buttenfield, B. P. (2000). Guidelines for the display of attribute certainty. *Cartography and Geographic Information Science*, 27(1):3–14.

[38] Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.

[39] Luo, A., Kao, D., and Pang, A. (2003). Visualizing spatial distribution data sets. In *VisSym*, volume 3, pages 29–38.

[40] MacEachren, A. M., Roth, R. E., O'Brien, J., Li, B., Swingley, D., and Gahegan, M. (2012). Visual semiotics & uncertainty visualization: An empirical study. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2496–2505.

[41] Messner, P. (2000). *Time shapes: a visualization for temporal uncertainty in planning*. Citeseer.

[42] Miksch, S. and Aigner, W. (2014). A matter of time: Applying a data–users–tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290.

[43] Munzner, T. (2009). A nested model for visualization design and validation. *IEEE transactions on visualization and computer graphics*, 15(6).

[44] Naumann, A., Hurtienne, J., Israel, J. H., Mohs, C., Kindsmüller, M. C., Meyer, H. A., and Hußlein, S. (2007). Intuitive use of user interfaces: defining a vague concept. In *International Conference on Engineering Psychology and Cognitive Ergonomics*, pages 128–136. Springer.

[45] Olston, C. and Mackinlay, J. D. (2002). Visualizing data with bounded uncertainty. In *INFOVIS 2002. IEEE Symposium on Information Visualization*, pages 37–40. IEEE.

[46] Pebesma, E. J., de Jong, K., and Briggs, D. (2007). Interactive visualization of uncertain spatial and spatio-temporal data under different scenarios: an air quality example. *International Journal of Geographical Information Science*, 21(5):515–527.

[47] Potter, K., Kirby, M., Xiu, D., and Johnson, C. R. (2012). Interactive visualization of probability and cumulative density functions. *International journal for uncertainty quantification*, 2(4).

[48] Potter, K., Kniss, J., Riesenfeld, R., and Johnson, C. R. (2010). Visualizing summary statistics and uncertainty. In *Computer Graphics Forum*, volume 29, pages 823–832. Wiley Online Library.

[49] Rit, J.-F. (1986). Propagating temporal constraints for scheduling. In *AAAI*, volume 86, pages 383–388.

[50] Robertson, G., Fernandez, R., Fisher, D., Lee, B., and Stasko, J. (2008). Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, 14(6):1325–1332.

[51] Sanyal, J., Zhang, S., Bhattacharya, G., Amburn, P., and Moorhead, R. (2009). A user study to compare four uncertainty visualization methods for 1d and 2d datasets. *IEEE transactions on visualization and computer graphics*, 15(6):1209–1218.

[52] Schweizer, D. M. and Goodchild, M. F. (1992). Data quality and choropleth maps: An experiment with the use of color. In *Proceedings of GIS/LIS*, pages 686–686. American Society for Photogrammetry and Remote Sensing.

[53] Shrestha, A., Zhu, Y., and Miller, B. (2014). Visualizing uncertainty in spatio-temporal data. In *ACM SIGKDD Workshop on Interactive Data Exploration and Analytics (IDEA)*.

[54] Slocum, T. A., Cliburn, D. C., Feddema, J. J., and Miller, J. R. (2003). Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. *Cartography and Geographic Information Science*, 30(4):299–317.

[55] Spear, M. E. and Spear, M. E. (1952). *Charting statistics*. McGraw-Hill.

[56] Tory, M. and Moller, T. (2004). Human factors in visualization research. *IEEE transactions on visualization and computer graphics*, 10(1):72–84.

[57] Tufte, E. R. and Schmieg, G. M. (1985). The visual display of quantitative information. *American Journal of Physics*, 53(11):1117–1118.

[58] Tukey, J. W. (1977). *Exploratory data analysis*. Reading, Mass.

[59] Viard, T., Caumon, G., and Levy, B. (2011). Adjacent versus coincident representations of geospatial uncertainty: Which promote better decisions? *Computers & Geosciences*, 37(4):511–520.

[60] Walny, J., Carpendale, S., Riche, N. H., Venolia, G., and Fawcett, P. (2011). Visual thinking in action: Visualizations as used on whiteboards. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):2508–2517.

[61] Walny, J., Huron, S., and Carpendale, S. (2015). An exploratory study of data sketching for visual representation. *Computer Graphics Forum*, 34(3):231–240.