University of Zurich[UZH]

Visualizing and communicating uncertainty for map-based decision-making: The case of uncertainty depiction in debris flow predictions

GEO 511 Master's Thesis

Author
Jana Bracher
16-725-269

Supervised by
Prof. Dr. Sara Irina Fabrikant

Faculty representative
Prof. Dr. Sara Irina Fabrikant

07.06.2022
Department of Geography, University of Zurich

Abstract

Any type of data is subject to uncertainty in one way or another. The prediction of natural hazards such as debris flows is no exception to this rule, especially in the face of ongoing climate change. Since maps are a valuable tool to depict scientific results, the visualization of uncertainty has occupied cartographers and visualization experts over the past decades. In this research, a large variety of different uncertainty visualization methods have been developed. However, testing their effectiveness and their impact on the decision-making process has not been on the forefront of research.

Therefore, the study at hand aimed at testing two types of uncertainty visualization methods (single-hue and multi-hue colour scheme; within-group variable) as well as two ways of communicating uncertainty in the map legend (numerical and verbal expressions; between-group variable) in debris flow prediction maps. A key aspect investigated in this study are the strategies applied to make decisions based on uncertain information. Additionally, the study makes use of eye tracking technology to infer on cognitive processes. Two research questions investigated the influence of the uncertainty visualization and communication methods on decision outcome, response time and decision-making strategy. The goal of the last research question was to gain insight into the sources of information which guide decision-making with uncertainty.

The empirical study showed that decision outcomes slightly varied between the two visualization methods. Additionally, the decision-making process seemed to be more complicated when uncertainty was communicated through verbal expressions, as shown by the significant difference in response time. Lastly, it was found that decisions were strongly guided by heuristics related to the uncertainty information as well as the distance parameter. Furthermore, a boundary effect, already observed in other uncertainty visualization studies, occurred. Most importantly however, the results indicate that the non-expert audience had trouble correctly interpreting the uncertainty information. Consequently, it is argued that map design choices might be of secondary importance as long as profound understanding of the concept of uncertainty is lacking among map readers. The study thus calls for more profound training of the public on the concept of uncertainty, its visualization in maps and ways to incorporate it into spatial decision-making.

Keywords: uncertainty visualization, uncertainty communication, spatial decision-making with uncertainty, debris flow prediction.

Acknowledgements

I would like to thank of my supervisors Dr. Sara Irina Fabrikant and Dr. Ian Ruginski for accompanying me over the course of a year in completing this thesis. The interesting discussions and the valuable feedback were always highly appreciated and significantly shaped the thesis.

Another thank you goes to the RAMMS development team at WSL, which provided me with a student license and technical support for the debris flow modelling. Additionally, I would like to thank all the debris flow experts and natural hazard practitioners for taking their time to have an interview with me.

I am also extremely grateful for all the participants of my study, who contributed a very valuable part to this thesis. Conducting the study would not have been possible without their help.

A thankful greeting to my friends who volunteered to proof read my thesis. I also want to thank you all for your support during our studies, for all the hours spent together in lectures and all our adventures outside of university so far.

Certainly the biggest thank you goes out to my family and my partner for supporting me not just during the Master's thesis but all throughout my studies. Without your unconditional support, I definitely would not be where I am now. Thank you so, so much!

Contents

List of Figures

List of Tables

Abbreviations

AOI	Area of interest
ARTool	Aligned Rank Transformation tool
DEM	Digital elevation model
GIS	Geographic information system / science
GVIS	Geographic visualization
H_0	Null hypothesis
H_A	Alternative hypothesis
HCI	Human-computer interaction
IPCC	Intergovernmental Panel on Climate Change
LMM	Linear mixed-effects model
RAMMS	Rapid Mass Movement Simulation
RQ	Research question
SD	Standard deviation
SDTS	Spatial Data Transfer Standard
TOI	Time of interest
$UVis^3$	Uncertainty Visualization cube

1. Introduction

1.1. Motivation

> *"(…) there is growing recognition, at least within the academic community, that: (1) the functionality of GIS needs to be enhanced to include ways of representing uncertainty; and (2) such representations need to communicate the uncertainty in a manner that is unambiguous, fully informative, and better able to facilitate decision-making."*
>
> *Hope and Hunter (2007)*

Uncertainty is an inevitable part of geographic data and models as, quite frankly, not everything is knowable (Harrower, 2003; Ruginski et al., 2016). Despite its omnipresence, there is no general consensus on how uncertainty is defined or how it should be visualized cartographically (Pang, 2008). Sources of uncertainty range from the data collection, over its processing, to the visualization of the data itself (Pang, 2001). Unfortunately, uncertainty information is frequently excluded from visualizations as the data is assumed to be completely correct (Brodlie et al., 2012). Consequently, map depictions convey a false validity (Clapham, 1992). It has been argued that uncertainty information should be included in map displays for the past decades. The reason is its potential to enable a clearer understanding of the displayed data and a higher confidence in decisions based upon it (Leitner and Buttenfield, 2000; MacEachren et al., 2005).

A variety of methods to visualize uncertainty have since been developed. However, empirical research investigating the effectiveness of said methods and the overall impact of visualizing uncertainty on decision-making remains inconclusive (Korporaal et al., 2020; MacEachren et al., 2005). In their systematic review of studies relating to uncertainty visualization, Kinkeldey et al. (2014) found a rising, nevertheless, small number of studies which went beyond categorizing types of uncertainties or suggesting new visualization methods. So far, uncertainty visualizations have been tested in different application fields such as infrastructure siting (Hope and Hunter, 2007; Leitner and Buttenfield, 2000), land cover classification (Drecki, 2002) as well as natural hazard prediction or weather forecasting for wildfires (Cheong et al., 2016), tornadoes (Ash et al., 2014; Klockow-McClain et al., 2020) and hurricanes (Cox et al., 2013; Millet et al., 2020; Ruginski et al., 2016). The most commonly known form of an uncertainty visualization is likely the so-called cone of uncertainty, developed by the National Hurricane Centre of the United States. The cone represents the potential trajectories of a hurricane and spans over areas which are expected to be affected by the event with a two-thirds probability based on data of the previous five years (Cox et al., 2013; Millet et al., 2020; Witt and Clegg, 2021).

Debris flows are a natural hazard common to Switzerland and entail high potential costs to public safety and infrastructure. Yet, their prediction is characterized by vast uncertainties. They are a gravitational process during which a mixture of water and sediment travel along a channel at velocities of up to 10m/s (Hirschberg et al., 2021a; Raetzo et al., 2002). The prediction of debris flows is frequently performed through modelling. However, since major knowledge gaps exist on appropriate debris flow parameters (Schraml et al., 2015), the prediction of future debris flow events is challenging and subject to major uncertainties. Additional uncertainty in predicting debris flows is introduced by climate change. The discussion and depiction of the uncertainty tied to those processes is not yet common in natural hazard management (Kubicek et al., 2012). This is also the case for the hazard maps created in Switzerland, in which different danger zones are separated by solid, deterministic borders (Trau and Hurni, 2007). The reason for the lack of uncertainty depiction in natural hazard visualizations can be traced back to missing cartographic guidelines on how to do so (Kunz et al., 2011a). Additionally, insecurities on the experts' side regarding how to express, let alone visualize uncertainties, have been found to prevent the visualization of uncertainty information (Fischhoff, 2012).

1.2. Goal & Overview of the Study

Although a variety of uncertainty visualization and communication methods have been proposed over the past decades, systematic, empirical research on the effectiveness of these methods is still sparse (Kinkeldey et al., 2017). Additionally, the visualization of uncertainty information depends on the specific field of application (Boukhelifa and Duke, 2009). Major research gaps exist in empirically gained knowledge on uncertainty visualizations for domain-specific applications. The study at hand thus aims at testing different uncertainty visualization and communication methods in the field of debris flow prediction mapping. Additionally, the study investigates the decision-making process involved in map-based processing of uncertainty information.

The user study is set up as a mixed design. Two uncertainty visualization methods – namely a single-hue and a multi-hue colour scheme – as well as two uncertainty communication methods – numerical and verbal expressions – are implemented. The visualization methods are based on the visual variables of colour value and colour hue. They represent a within-group variable, while the communication methods vary between groups. The task posed to the non-expert participants is to judge the potential damage a debris flow could cause at a specific house location indicated on a map display. The debris flow uncertainty information on the map display shows the spatial uncertainty of a location being reached by the flow. Participants choose from seven predefined damage categories represented by a Likert scale to estimate the potential damage.

This study will help to increase the amount of knowledge generated by empirical studies in uncertainty visualization, as called for by MacEachren et al. (2005). The methods employed in this study will also support a general shift in uncertainty visualization studies from merely reporting which method performed best towards investigating why certain effects were observed (Hullman et al., 2019). To the knowledge of the author, this is the first empirical study on uncertainty visualization specifically addressing the phenomenon of debris flows. Additionally, no previous study comparing verbal and numerical uncertainty communication in a map-based context was found. It thus represents a first attempt at uncertainty visualization and communication for this highly relevant natural process. The results can inspire natural hazard practitioners and debris flow experts to depict uncertainty in their visualizations in an aim to put the inherent uncertainty of debris flows into the spotlight.

The research questions as well as the expected results of this study are presented below. The next chapter of the thesis investigates the state-of-the-art in uncertainty research, how uncertainty can be visualized cartographically and how it can influence human decision-making. Chapter 3 contains detailed information on the applied methods. The results of the study are presented in Chapter 4 and discussed in Chapter 5. Lastly, conclusions are drawn from the study and an outlook on potential follow-up research is provided.

1.3. Research Questions

Based on the existing research gaps briefly mentioned above, the following three research questions are posed for this thesis:

RQ1: How do different visualization methods (single-hue, multi-hue) for displaying uncertainty in debris flow predictions influence the decisions (decision outcome, response time, decision-making strategies) of map readers?

RQ2: Is there a difference in decision-making if subjects are presented with verbal versus numerical expressions to communicate the uncertainty information?

RQ3: Are decisions guided by the uncertainty information or by additional information on the map display (slope and distance from riverbed)? What role do the slope and distance information play in the decision-making strategies?

1.3.1. Hypotheses & Expected Results

Based on the literature review conducted in advance of the study, the expected results are framed below. Note that expectations regarding the decision-making strategies are discussed across all three research questions in a separate subchapter.

Generally speaking, no correlation between decision outcomes (expressed in the form of damage estimates) and the uncertainty value is expected due to the nature of the visualized uncertainty information. Similar damages can be expected independent of whether the uncertainty of a house location being affected is high or low. However, since it is known that humans struggle to understand uncertainty on a theoretical level (Joslyn and Savelli, 2021), it is possible that participants misinterpret the information. This could mean that damages are judged to be higher if uncertainty is low.

1.3.1.1. RQ1: Uncertainty Visualization

The hypothesis regarding research question 1 is that the two intrinsic visualization methods – single-hue (colour value) and multi-hue (colour hue) – lead to a difference in the damage estimate but not the response time.

The first part of the hypothesis is supported by results of various studies such as Ash et al. (2014), Klockow-McClain et al. (2020), Kübler et al. (2020) and Leitner and Buttenfield (2000), who each found differences between various types of intrinsic visualization methods. Unfortunately, studies directly comparing colour value and colour hue are rather sparse. Cheong et al. (2016) found that colour hue led to more correct answers when using it to depict uncertainty in tornado forecasts. However, this difference was not significant. Therefore, it is expected that the difference in damage estimate in this study might only be minor. The following is expected in terms of the directionality of this difference: Damage estimates using the single-hue colour scheme could be higher as the method is uncertainty evoking. Thus, the regions of low uncertainty are more salient. This is, however, only true if the participants struggle to interpret the uncertainty information as mentioned above.

Cheong et al. (2016) did not find significant differences in the response time when using colour value and colour hue. Since both methods tested in this study are based on characteristics of colour, it is assumed that they are processed similarly. Consequently, no difference in the response time is expected.

Lastly, multiple studies such as Cheong et al. (2016), Klockow-McClain et al. (2020) and Miran et al. (2019) found that participants preferred the colour hue method for decision-making with uncertainty over others. Similar results could be found in the preference task of the study at hand.

1.3.1.2. RQ2: Uncertainty Communication

The hypothesis of the second research question is that differences in the damage estimate and the response time can be observed between the two uncertainty communication groups.

No study has directly compared numerical and verbal uncertainty expressions in a map-based study to the knowledge of the author. This is why the hypothesis can only be inspired by results of text-based studies. If participants interpret the numerical and verbal uncertainty similarly, no major differences are expected. However, if Budescu et al. (2012, 2009) are right that verbal expressions are not correctly matched to numerical values, then a difference in the decision outcome could be found. Yet, again the directionality of this difference is not clear.

Doyle et al. (2014) found that participants presented with verbal uncertainty expressions first converted these into a numerical estimate before making their decision. This could also be the case in this study and would lead to a higher response time for the verbal uncertainty group compared to the numerical one.

Inspired by various studies mentioned in Chapter 2.6, it is expected that the communication method has an influence on the trust rating and the task difficulty rating. Trust is expected to be higher for the numerical uncertainty group as these expressions are more precise (Wallsten and Budescu, 1995). Since the audience tested in this study is non-expert, the task difficulty rating is expected to be lower for the verbal group compared to the numerical one due to their vagueness and the natural language used in verbal expressions (Wallsten and Budescu, 1995).

1.3.1.3. RQ3: Influence of Additional Information

For the third research question it is expected that the uncertainty information has the largest influence on the damage estimate. Nevertheless, the distance between the house location and the riverbed as well as the slope information are also expected to be considered during the decision-making process. Between the two, the expectation is that the distance is more important than the slope. As explained above, damage estimates are not expected to correlate with the uncertainty information. Damage estimates could increase the higher the slope and the shorter the distance are due to the physical properties of debris flows.

The hierarchy in influence between the uncertainty and the additional information on the decision-making process could be guided by the varying task-relevance of these sources of information. A strong influence of the distance has also been found in uncertainty visualization studies such as Klockow-McClain et al. (2020) and Ruginski et al. (2016). Another effect relating to the distance, which will likely occur in the results, is a boundary effect. This effect could lead to significantly lower damage estimates for house locations right outside the debris flow shape, compared to ones right inside of it. Consistent evidence of this effect has been found in various studies (e.g., Ash et al., 2014; Klockow-McClain et al., 2020; Ruginski et al., 2016) in the context of the hurricane cone of uncertainty and tornado forecasting. In terms of the strength of the effect, it is expected to be stronger for the multi-hue stimuli. This could be caused by the perceived fuzziness of the single-hue method and reinforce results found with a fuzzy hurricane cone tested by Ruginski et al. (2016).

1.3.1.4. Decision-Making Strategies

With the applied design of the study, decision-making strategies can be described overall and potential differences caused by the two uncertainty communication methods can be uncovered. However, no distinct differences caused by the two visualization methods can be systematically detected. Nevertheless, building up on the expectations of RQ3, it is expected that decision-making strategies will strongly focus on the uncertainty information represented through the colour schemes and the map legends. Yet, the additional information on the slope as well as distances is also expected to be taken into account.

In terms of the differences caused by the communication methods, it is possible that the verbal groups is more likely to use the additional information to make their decision compared to the numerical group. This could be caused by the vagueness of the verbal uncertainty expressions (Joslyn and Savelli, 2021).

Differences due to the visualization methods being mentioned in the post-test questionnaire could be related to the preference on the two colour schemes and have an effect on the perceived difficulty of the task. Therefore, participants could mention a lower task difficulty for the multi-hue colour scheme as it is expected that this method is preferred.

2. Background

2.1. Geographic Visualization

> *"(…) the multidisciplinary nature of the visualization researcher in combining engineering, science, and art. Engineering in that the visualizations are problem driven with users trying to understand or look for features in their data sets and the visualization researchers specifying the best practice approach. Science in that the visualization researcher also needs to draw upon various established fields such as perceptual and cognitive psychology, mathematical and physical analyses, etc. Visualization is also an art in that the results need to be tailored to the particular task, needs, and occasion."*
>
> *Pang (2008)*

Geographic Visualization (GVIS) forms a bridge between the fields of cartography, with its rich repository of mapping technology, and scientific visualization (Clapham, 1992; Howard and MacEachren, 1996). Through GVIS, map makers have the ability to reduce the vast amount of available information, convey it in a quickly readable form and ultimately enable data exploration (Borland and Taylor, 2007; MacEachren, 1995). Geographic visualizations – or maps – have found their way into the most diverse fields of application due to their value for communicating information effectively (Chesneau et al., 2005).

Buttenfield and Ganter (1990) distinguish three different purposes for which a visualization can be used: analysis/inference, illustration and decision-making. A visualization for analysis/inference serves as a way to explore data, elaborate a hypothesis and investigate errors within a dataset. The visualization herein represents a tool which is only in use during data exploration. This is also why these visualizations can seem abstract from a design perspective. In comparison, visualizations for illustration are seen as a rendered form to present a reality by implementing design principles. If a visualization is created with the purpose to serve for decision-making, it is most important that a summary of the information is presented in a way "so that decision-making is neither obstructed nor biased" (Buttenfield and Ganter, 1990). Knowledge on the decision-makers as well as the process of decision-making (see Chapter 2.5) are key aspects when creating these types of visualizations.

According to Quispel and Maes (2014), visualizations aim to fulfil the criteria of accuracy, efficiency, attractiveness and aesthetics. The latter two are of special importance when moving beyond scientific visualization towards a more diverse audience. This is directly related to the illustration visualizations presented above. However, despite the conquest of geographic visualization, it must be noted that the assumption that map readers of various backgrounds are capable of correctly interpreting visualizations, may not always be valid (Fabrikant et al., 2010).

2.2. Visual Processing of Geographic Visualizations

While cartography and its well-established principles play a key role in geographic visualization, another relevant aspect is how map users view and process a visualization. Therefore, the most important theories of the field of cognition and visual scene processing will be briefly presented.

The main goal of discussing cognitive theories in the context of visualization is to improve cartographic methods and enable predictions on how map readers perceive a visualization (MacEachren, 1995). Cognition can be seen as the entirety of mental processes which constitute to thinking. These include perception, memory, reasoning and attention (Peterson, 1994; Van Der Bles et al., 2019). Narrowing this down to the field of visualization, visual cognition is defined as the process of "deriving meaning from external representations of visual information" (Padilla et al., 2018).

Visual information processing can be conceptualized into a simplistic model with three types of memory stores (see Figure 1). These stores consist of the iconic memory, the short-term visual store as well as the long-term visual memory (Peterson, 1994). The iconic memory recognizes sensory information such as an object on a map. Attention is necessary to move from the iconic memory to the short-term visual store. To focus on the perceived object and interpret it in the short-term visual store, long-term visual memory is applied on the information. Overall, the processing of visual information can be seen as a process of pattern recognition (Peterson, 1994).

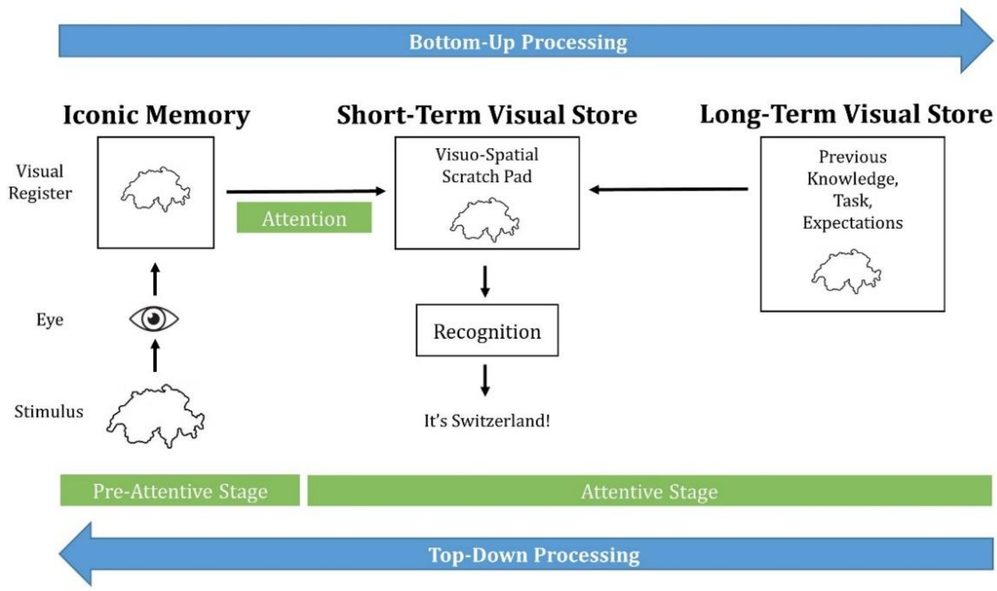


Figure 1: Concept of visual processing (figure designed by author, inspired by Peterson (1994), extended with information from Harold et al. (2016), Padilla et al. (2018) and Wolfe and Horowitz

Due to the anatomy of the human eye, processing at high-resolution is restricted to the fovea of the eye (Wolfe and Horowitz, 2004). Therefore, only a small subset of a visualization can be attended at a time. The capacity of attention is limited, which is why objects in a visualization compete for visual attention (Hegarty et al., 2012). Bottom-up and top-down processes guide the trajectory visual attention. Processes in bottom-up direction are task-independent (Itti et al., 1998). Likely the most important determinator from a bottom-up perspective are attention-guiding properties of symbolization, which are processed pre-attentively (Wolfe and Horowitz, 2004). Examples of such properties are some of the visual variables originally developed by Bertin (1983). More information on the visual variables, their development and attention-guiding properties are discussed in Chapter 2.4.1. After attention is drawn to a specific region, the attentive stage is entered during which processing occurs consciously (Krassanakis, 2013). Therefore, visual attention can determine whether some information contained in a map is processed or overlooked (Harold et al., 2016). From a top-down perspective, processing includes the recognition and interpretation of patterns (Padilla et al., 2018). This type of processing is guided by previous knowledge or a specific task (Harold et al., 2016; McMahon et al., 2015).

To summarize, visual cognition is composed of a system with three types of memory stores (Peterson, 1994). Visual processing includes processing in bottom-up direction – such as attention guidance – as well as top-down direction – for instance previous knowledge (Itti et al., 1998; Padilla et al., 2018). The processes in these two directions intertwine. Thus, by perceptually processing a map display using visual variables, map readers are simultaneously understanding it cognitively (Roth, 2017).

2.3. Uncertainty

2.3.1. Definition of Uncertainty

As mentioned at the very beginning of this thesis, uncertainty is an inevitable part of any type of data whether it was collected through measurements or created through modelling (Gong and Chen, 1992; Ruginski et al., 2016). This is especially true when predictions are made about future developments (Pappenberger and Beven, 2006; Zhang and Goodchild, 2002). Apart from this, research on uncertainty is challenged by the vast variety of definitions of uncertainty circulating in different fields, none of which serve as a standard definition (Pang, 2008). However, a clear definition is a strong prerequisite for discussing and visualizing uncertainty information in cartography and other disciplines (Buttenfield, 1993; Hullman et al., 2019).

As such uncertainty is an overarching concept including terms such as error, vagueness and imprecision (Kinkeldey et al., 2017). Thomson et al. (2005) defined uncertainty as "the degree to which the lack of knowledge about the amount of error is responsible for hesitancy in accepting results and observations without caution". Therefore, the key component of uncertainty is the absence of accurate knowledge about a topic. Uncertainty could then be quantified as the amount of knowledge necessary to arrive at the truth (Zhang and Goodchild, 2002). In the context of modelling events in the future, Joslyn and Savelli (2021) defined uncertainty as the likelihood of a specific event taking place. One concept closely related is data quality, which is sometimes seen as a synonym for uncertainty (Boukhelifa and Duke, 2009). The two concepts are inversely related, meaning that a dataset with low data quality contains major uncertainty (Pang et al., 1997).

2.3.2. Types of Uncertainty

The types of uncertainty between which one can distinguish are just as diverse as its definitions. One common distinction for uncertainty in model outputs is made between aleatory and epistemic uncertainty. While aleatory uncertainty describes the natural uncertainty which is inherent to a phenomenon, the epistemic uncertainty is caused by gaps in the knowledge on a process (Kunz et al., 2011b; Van Der Bles et al., 2019). Padilla et al. (2021a) mention another type of uncertainty in this context: The ontological uncertainty gives insight into the deviation between a natural process in reality and its representation through a model.

Another distinction can be made between direct and indirect uncertainty. Direct uncertainty is related to a specific fact or a number and can be quantified. This type of uncertainty is usually expressed through confidence intervals or probabilities. Indirect uncertainty describes the quality of the knowledge upon which facts and numbers are based. Its quantification is not directly possible and it can only be indicated in the form of qualitative confidence (Padilla et al., 2021b; Van Der Bles et al., 2019).

2.3.3. Uncertainty Typologies

Due to the variety of definitions and types of uncertainty circulating in literature, typologies have been developed to organize the chaos (Zhang and Goodchild, 2002). The most important aspect of a harmonized typology is that it includes all types of uncertainty relevant to a specific application. These terms must then be applied consistently. If this is the case, a typology can provide guidance in investigating and communicating data uncertainty (Thompson and Warmink, 2016). The presence of different categories of uncertainty in typologies make it evident that a specific data set can contain more than one type of uncertainty (MacEachren et al., 2012).

The first typology presented here was developed in the context of the Spatial Data Transfer Standard (SDTS) (USGS, 1997). The aim of this standard is to create guidelines on how to transfer data as well as its metadata in a consistent way. One part of the standard specifically defines the five categories of spatial data quality: lineage, positional accuracy, attribute accuracy, logical consistency and completeness (see Table 1).

Table 1: Uncertainty typology as proposed in the SDTS (all definitions are quotes from PlanGraphics Inc. (1996)).

Category	**Definition**
Lineage	"Information on sources, update activity with dates, and processing steps that have transformed the data."
Positional Accuracy	"Information about how closely coordinate values of map features match their true location. (…)"
Attribute Accuracy	"Information on the error in the values of attribute data elements included in a transfer. (…)"
Logical Consistency	"An indication of the graphic quality and topological integrity of a digital map."
Completeness	"Information about selection criteria for inclusion of map features, minimum thresholds in map compilation (…), and the exhaustiveness of features mapped."

The two categories which are especially relevant for the study at hand are the positional and the attribute accuracy. The positional accuracy describes a spatial error in the location of data whether that is a specific point or a boundary (Gong and Chen, 1992; Kinkeldey et al., 2017). Attribute accuracy does not relate to the spatial properties of data (Gong and Chen, 1992). Depending on the type of data, the interpretation of attribute accuracy is twofold. On the one hand, it describes the accuracy of the measurement of a specific, continuous attribute. On the other hand, it represents the accuracy of the classification of categorical attributes (Thomson et al., 2005).

Table 2: Uncertainty typology developed by Thomson et al. (2005) for geographical information (all definitions are quotes from Thomson et al. (2005)).

Category	**Definition**
Accuracy/error	"Difference between observation and reality"
Precision	"Exactness of measurement"
Completeness	"Extent to which info is comprehensive"
Consistency	"Extent to which info components agree"
Lineage	"Conduit through which info passed"
Currency/timing	"Temporal gaps between occurrence, info collection & use"
Credibility	"Reliability of info source"
Subjectivity	"Amount of interpretation or judgment included"
Interrelatedness	"Source independence from other information"

Another typology of uncertainty specifically developed for geographical information, which builds up from the one just presented, was conceptualized by Thomson et al. (2005). The overall goal of the typology is to establish guidance in the visualization of these different categories of uncertainties. In this typology, a total of nine uncertainty categories are defined: accuracy/error, precision, completeness, consistency, lineage, currency/timing, credibility, subjectivity and interrelatedness (see Table 2). Each of these categories can be applied to the attributes as well as the spatial and temporal aspects of geographical data.

2.3.4. Sources of Uncertainty

Generally speaking, uncertainty occurs because perfect knowledge on a phenomenon or process is not attainable (Harrower, 2003). Pang et al. (1997) suggested that uncertainty can be introduced at any point of a data visualization process ranging from the data collection, its processing and transformation to the visualizing itself. During the data acquisition, uncertainty can be caused by inaccurate measurements (Boukhelifa and Duke, 2009). Specifically for data generated through running numerical models, which simulate a certain process, uncertainty is associated with the parameters fed into the model. The fact that the model is essentially a simplification of the real process introduces additional uncertainty (Kunz et al., 2011b; Pang, 2008). Uncertainty is propagated when data is aggregated during its processing (Buttenfield and Ganter, 1990). Additional uncertainty can be introduced during smoothing, filtering or sub-sampling procedures (Pang, 2008). It can also be added through generalization during the visualization of data (Kunz et al., 2011b). Lastly, it can be introduced after a visualization is created due to misinterpretation by the map reader (Pang, 2008). It is thus clear that the uncertainty present in data can stem from a variety of sources.

2.3.5. Uncertainty in Geographic Information

Since geographic information represents an abstraction of reality, all spatial data is subject to uncertainty (Schweizer and Goodchild, 1992). Therefore, it does not come as a surprise that uncertainty varies across space just as the data itself does (Zhang and Goodchild, 2002). MacEachren et al. (1998) warned that the neglection of this uncertainty information could lead to false interpretations of the presented data patterns. Roth (2009) identified three challenges when dealing with uncertainty in geographic information:

1. "determining the current involvement of uncertainty at different stages in the geographic information life cycle"
2. "identifying the many forms that uncertainty can take in this process"
3. "understanding the influence these forms have on the use of geographic information"

These challenges make clear that uncertainty as a type of *data about data* must not be neglected. Therefore, a strategy on how to manage uncertainty in geographic information has long been called for by Hunter (2005). He thus proposed that the data quality must be closely monitored when producing data. Additionally, software developers should create the necessary tools to document and communicate the uncertainty information. Subsequent to managing uncertainty in geographic information, it is crucial that the uncertainty associated with it is also visualized in order to use geographic information in a transparent way (Roth, 2009).

2.4. Uncertainty Visualization

Unfortunately, cartographic visualization methods were mostly designed under the assumption that the represented data does not contain any uncertainty (Brodlie et al., 2012). One aspect questioned by Leitner and Buttenfield in 2000 was whether established cartographic methods could be similarly applied on uncertainty information. This question is still up for debate. Consequently, uncertainty of scientific results or modelling analyses is hardly ever communicated or displayed to their audience (Padilla et al., 2015; Pappenberger and Beven, 2006). Nevertheless, the visualization of uncertainty in spatial data is especially important if said data is used for decision-making (Drecki, 2002). It has the potential to lead to more informed decision-making and more profound understanding of the reality by uncovering the patterns of spatial uncertainty (Zhang and Goodchild, 2002). The ultimate goal of uncertainty visualization is to help map readers interpret and incorporate uncertainty information when using data for decision-making (Pang et al., 1997).

One challenge associated with the visualization of uncertainty lies in the fact that adding uncertainty information to a display can make things more complicated. This does not just affect the handling of the data but also the cognitive efforts to process the uncertainty information (Boukhelifa and Duke, 2009; Pang, 2008). Methods to visualize uncertainty must thus aim at adding information without distracting from the data that is already present (Cedilnik and Rheingans, 2000). There is a fine line between choosing suitable uncertainty visualization methods and applying unsuitable ones. A method could be unsuitable for uncertainty visualization due to its complexity, the inappropriateness of the used metaphor or the flawed development of a colour scheme (Gershon, 1998). Through the application of unsuitable methods, the uncertainty associated the displayed data can further increase (MacEachren et al., 2005). Likely the largest challenge in uncertainty visualization is that knowledge on the effectiveness of the proposed methods contains major gaps (Aerts et al., 2003). It is thus proposed that research in uncertainty visualization, such as the study at hand, should aim at closing those gaps instead of proposing additional visualization methods. Lastly, past studies have shown that choosing an uncertainty visualization method can be highly application-dependent (Boukhelifa and Duke, 2009; Hunter, 2005; Joslyn and Savelli, 2021). This further challenges empirical research on uncertainty visualization. However, it can also serve as a guidance for systematic, application-specific research.

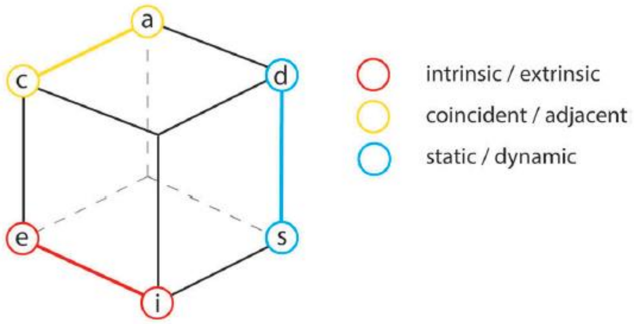


Figure 2: The Uncertainty Visualization cube (UVis³) – a framework for categorizing uncertainty visualization methods (Kinkeldey et al., 2014).

As mentioned above, an abundance of different uncertainty visualization methods has been developed during the past decades. In an aim to categorize these methods, Kinkeldey et al. (2014) developed the Uncertainty Visualization cube (UVis³) as displayed in Figure 2. The cube is composed of the three dichotomies: intrinsic/extrinsic, coincident/adjacent and static/dynamic. These three contrasts will be discussed separately and illustrated with studies in the following subchapters. Due to their relevance to the study at hand, intrinsic methods will be presented in more detail, while the other types of methods are only considered briefly to complete the picture.

2.4.1. Intrinsic Uncertainty Visualization

Intrinsic uncertainty visualization methods incorporate the uncertainty information on the data symbolization itself by making use of the so-called visual variables (Bisantz et al., 2009; Kinkeldey et al., 2014). The visual variables were first proposed by Bertin (1983) (originally published in French in 1967) and likely represent the most important concept in cartographic design. These variables "describe the graphic dimensions across which a map or other visualization can be varied to encode information" (Roth, 2017). Bertin (1983) defined the following seven visual variables: *location, size, (colour) value, texture, colour (hue), orientation and shape* (see Figure 3). Colour saturation was added to the list by Morrison (1974). Lastly, MacEachren (1992) developed the visual variable of focus specifically for depicting uncertainty. He later split this variable up into crispness/clarity, resolution and transparency (MacEachren, 1995).

When using intrinsic methods, cartographers encode uncertainty by manipulating visual variables (Kinkeldey et al., 2014). These can be characterized in terms of their properties of being associative and selective as well as in their ability to imply an order. Variations of an associative visual variable are interpreted as equal in weight and they can be grouped visually. If a visual variable is selective, one can attend and process variations of it individually and extract patterns quickly (Roth, 2017). These properties give insight into their suitability to encode different types of data (e.g., nominal, ordinal and numerical) (for an overview see Roth (2017)). Intrinsic methods are the most commonly applied type of method in uncertainty depiction (Kinkeldey et al., 2014) and the number of proposed methods in this realm is vast. A potential reason for this is that map makers are extremely familiar with the principle of visual variables. However, one drawback of intrinsic methods is that small changes in uncertainty can be hard to distinguish from one another (Kunz et al., 2011b).

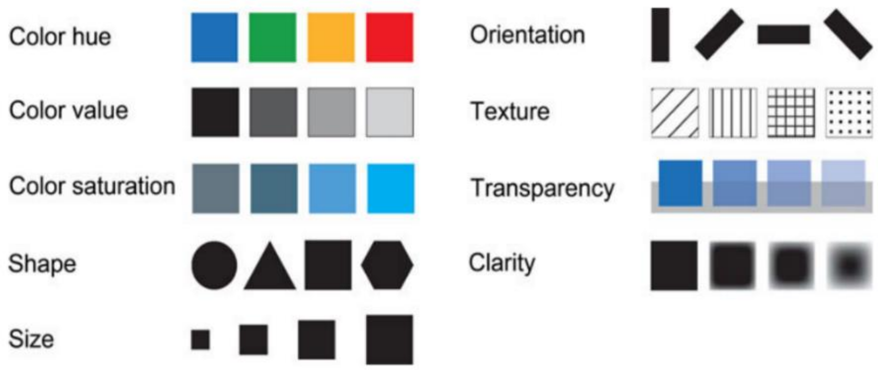


Figure 3: Overview of all visual variables with the exception of resolution (Kunz et al., 2011b).

One theory with a high impact on intrinsic uncertainty visualization methods is that of visual semiotics. It "proposes that features that viewers spontaneously interpret as conveying uncertainty will be more effective than features that do not evoke uncertainty associations" (Padilla et al., 2021a). This is again tied to the concept of attention and attention-guiding properties. The cognitive load of processing the visual information is lower when applying visual variables as proposed by the concept of attention-guiding properties, which enables faster decision-making (Swienty et al., 2008). The goal of visualizing uncertainty intrinsically is thus to evoke a feeling of uncertainty. A classic example of an uncertainty-evoking method is altering the visual variable of transparency and concealing uncertain data, as proposed by MacEachren (1992) (see Figure 4A). Therefore, this method makes use the spatial metaphor of fog. Transparency could, however, also be used in addition to another visual variable, such as colour value, to encode uncertainty redundantly (Zuk and Carpendale, 2006). Another method proposed by MacEachren (1992), which is especially suitable for linear features or polygon outlines, is fuzziness (see Figure 4B). A certain outline is represented as a sharp boundary, while more uncertain outlines are displayed fuzzily. This concept was successfully implemented for hurricane forecasts in a study by Millet et al. (2020). Another highly popular way to depict uncertainty intrinsically is by using some aspect of colour. These methods include the usage of a colour map, which guides the map reader in investigating the uncertainty of the data (Pang, 2008). Johannsen et al. (2018) have argued that colour value is more intuitive and uncertainty-evoking compared to the other two dimensions of colour – hue and saturation.

A - fog / transparency **B – clarity / fuzziness**

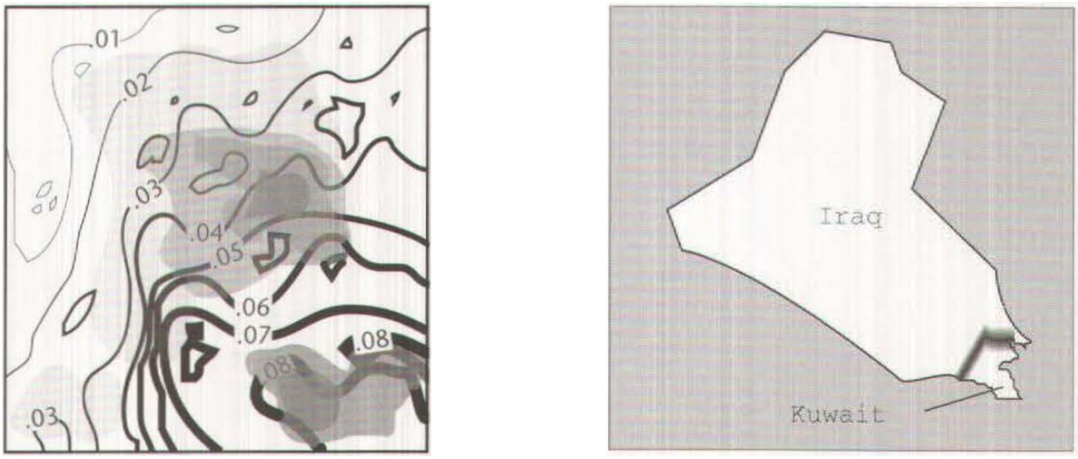


Figure 4: Intrinsic uncertainty visualization methods proposed by MacEachren (1992): A – fog method to display the uncertainty in ecological risk due to ozone. B – clarity method to depict the uncertain border between Iraq and Kuwait.

MacEachren (1992) saw the largest potential to display numerical uncertainty in the visual variables of size and colour value. However, this statement was of theoretical nature and not based on empirical results. MacEachren et al. (2012) later conducted an empirical study to explore the intuitiveness and suitability of different visual variables to display general uncertainty in detail (see Figure 5). They found that participants rated fuzziness, location and value as the most intuitive representations of uncertainty with mean ratings of around 6 on scale from 1 – *illogical* to 7 – *logical*. These were followed by arrangement, size, transparency, grain and colour saturation. All the remaining visual variables scored below the overall average of intuitiveness ratings. Although colour saturation was often referred to as a very valuable method to depict uncertainty (MacEachren, 1992), it was not among the highest ranks. A potential reason for this is that colours with low saturation, usually representing data of high uncertainty, tend to become grey and are thus hard to distinguish from one another (Drecki, 2002). Therefore, colour saturation is not considered to be a suitable uncertainty visualization method nowadays (Kinkeldey et al., 2014). It was also shown that the directionality of how a visual variable is implemented can have an influence on how intuitive participants judged the method to be (MacEachren et al., 2012). Transparency, for instance, performed better when high transparency represented high uncertainty. The contrary was proposed by MacEachren (1992) in the context of the fog method mentioned above.

A selection of further studies, which empirically tested different intrinsic methods, as well as their results are more thoroughly discussed in Chapter 2.4.5.

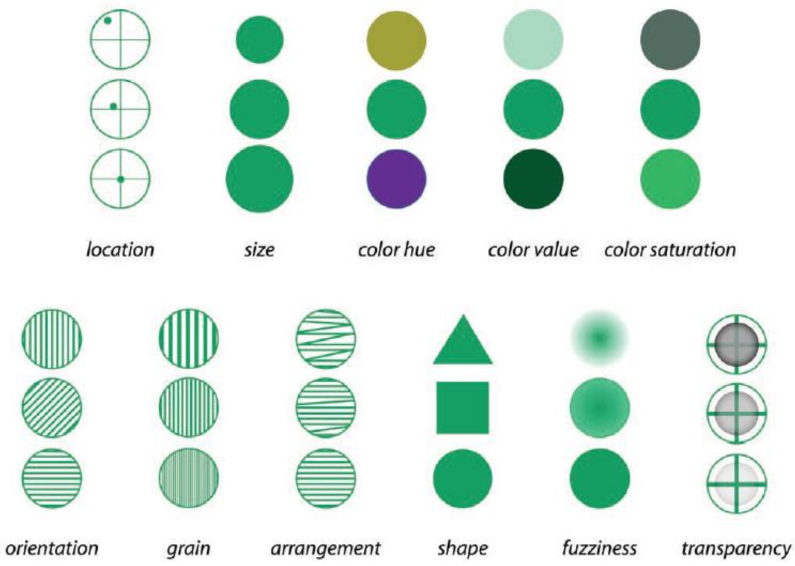


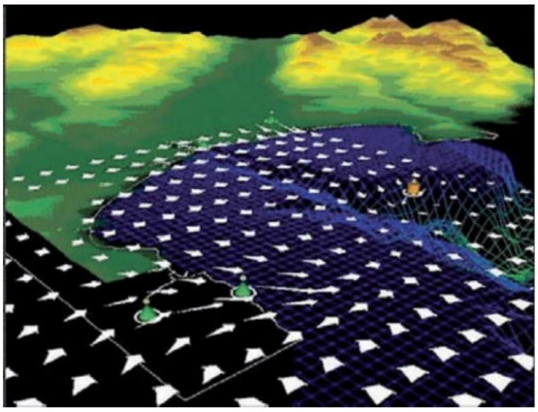
Figure 5: Visual variables tested on their intuitiveness to represent general uncertainty in the study by MacEachren et al. (2012). Each visual variable was tested in both directions (e.g., the purple hue once represented a high and once a low uncertainty).

2.4.2. Extrinsic Uncertainty Visualization

Extrinsic methods are characterized by adding geometry to represent uncertainty information (Pang et al., 1997). This has the advantage that the underlying data can be visualized unaltered. Glyphs are likely the most common approach of adding geometry. They were first applied to depict uncertainty in ocean currents and wind in the form of arrows by Wittenbrink et al. (1996) (see Figure 6A). The size, length or direction of the glyphs are distorted depending on the amount of uncertainty in the data (Potter et al., 2012). Other forms of glyphs include thermometers or error bars which are added to the map displays (Gershon, 1998). After developing the glyphs to depict uncertainty, Wittenbrink et al. (1996) conducted an empirical study to investigate whether map readers, in their case naval officers, were able to correctly interpret the uncertainty. They found that similar errors were made with the uncertainty glyphs and a deterministic arrow glyph during the decoding of the information. Since participants were able to read the uncertainty glyphs, they concluded that these glyphs represent an improvement compared to their deterministic counterpart due to their higher information content.

Hope and Hunter (2007) used glyphs in the form of bars to represent uncertainty in an airport siting study and found rather negative results. Contrary to their expectations, participants chose locations with a low uncertainty more frequently even if two locations were equally suitable. They concluded that participants made irrational choices when uncertainty was included in map displays in the form of glyphs. They traced these results back to difficulties with interpreting the uncertainty information.

A - glyphs



B1 – width **B2 - sharpness** **B3 – noise** **B4 – amplitude**

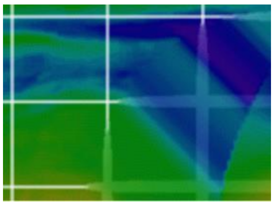
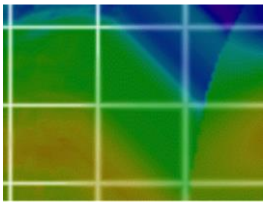
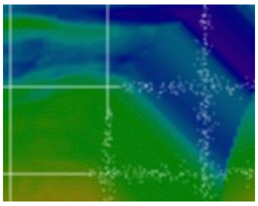
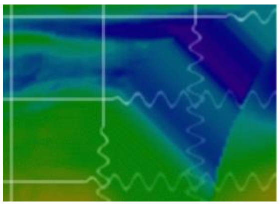





Figure 6: Extrinsic uncertainty visualization methods: A - glyphs as developed by Wittenbrink et al. (1996) (printed by Pang (2008)). B - different variations of the grid-based approach developed by Cedilnik and Rheingans (2000).

Another extrinsic approach is to overlay the data with contour lines to visualize uncertainty (Pang, 2008). The amount of uncertainty is then represented through varying the visual variable of the contour geometries such as the thickness (size) or the colour value. The grid-based visualization – a strongly related approach – was developed by Cedilnik and Rheingans (2000). To apply this method, the data is covered by an equally spaced grid. Again, the symbology of the grid is deformed depending on the amount of uncertainty in the data. They proposed different approaches to represent uncertainty in the grid lines: width/size, sharpness/fuzziness, noise and the amplitude of a wave (see Figure 6B1-4). The advantage of this method is that map readers are used to overlying grids in maps, so it does not strongly distract from the underlying data. A disadvantage mentioned by Cedilnik and Rheingans (2000) is the high computational effort of creating the visualizations. However, this is likely not the case anymore due to the technological advances.

During the development of a software to visualize uncertainty in the global water balance, Slocum et al. (2003) conducted an evaluation of the system through interviews. They found that experts preferred extrinsic methods because of the higher level of detail displayed in the visualizations. Decision-makers, on the other hand, preferred intrinsic visualizations as they helped them to gain a quick overview of the data. Therefore, the decision on whether to represent uncertainty intrinsically or extrinsically could be dictated by the target audience of the visualization.

2.4.3. Coincident vs. Adjacent Uncertainty Visualization

The methods presented so far included both data and its uncertainty in a single display. They thus represent coincident methods. Adjacent uncertainty visualization methods are composed of two map displays: one shows the thematic data, while the other one displays the uncertainty associated with the data (Kunz et al., 2011b). A drawback of adjacent displays is the number of eye movements necessary to investigate the maps as the two displays must be matched cognitively. This is not the case for coincident maps, where the data and the uncertainty can be processed simultaneously. However, coincident displays can be rather complicated and cluttered due to the high information content (Kinkeldey et al., 2014).

An interesting coincident approach developed by Correll et al. (2018) are so-called value suppressing uncertainty palettes (see Figure 7). These palettes include colour schemes which are more differentiated for data with low uncertainty. Data with very high uncertainty is displayed similarly no matter what the data value is.

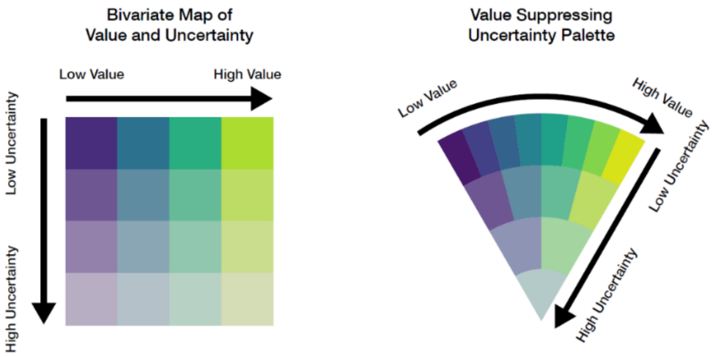
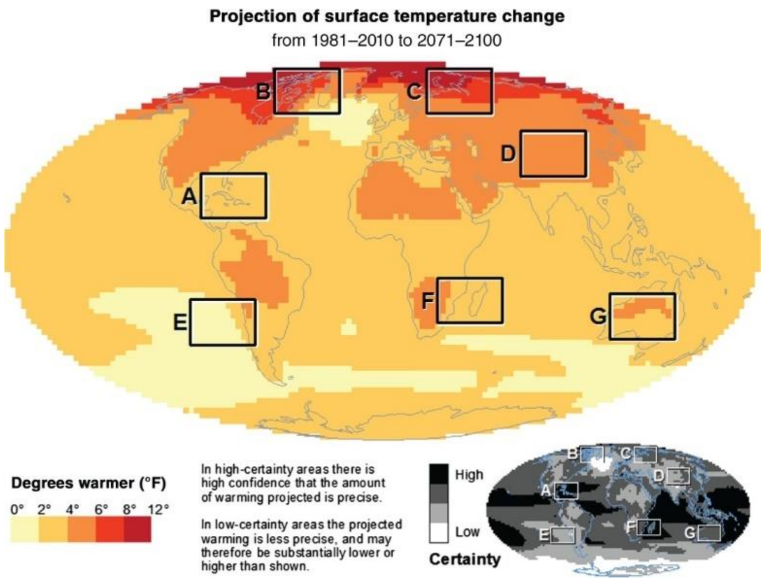


Figure 7: Coincident uncertainty visualization method in the form of a classic bivariate map and a value suppressing uncertainty palette (Correll et al., 2018).

Retchless and Brewer (2016) developed an adjacent as well as a variety of pattern- and colour-based coincident map displays to depict uncertainty associated with the surface temperature change caused by global warming (see Figure 8). The participants were asked to rank different regions on the map in terms of their temperature as well as their associated uncertainty. They found that the adjacent method led to the most accurate rankings for both the temperature and the uncertainty. While differences between the control group (adjacent method) and the different types of colour-based coincident maps were low for rating the temperature, the ranking of the uncertainty yielded significant differences. They also found that participants were most accurate with coincident maps when colour was used to represent the temperature and a pattern for uncertainty. However, they also noted that the choice of the visualization method depends strongly on the goal of the map maker.

A



B

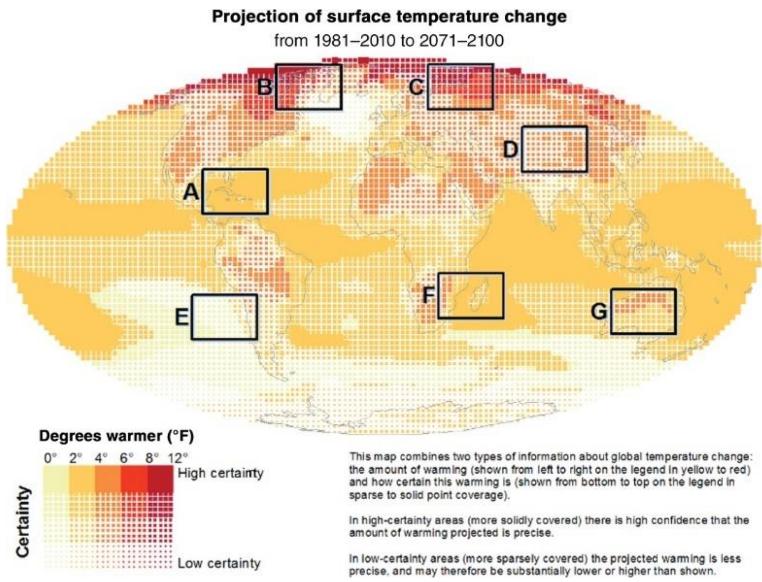


Figure 8: Adjacent (A) and coincident (B) uncertainty visualization of the projection of surface temperature change as tested by Retchless and Brewer (2016).

2.4.4. Static vs. Dynamic Uncertainty Visualization

All the uncertainty map displays shown thus far are static, which means that map readers do not have the possibility to interact with the display. Dynamic uncertainty visualizations make use of animated map displays (Pang, 2001). Parameters which can be used to represent uncertainty dynamically are the speed of motion, the range of motion or motion blurring (Pang et al., 1997). First methods to use animation in uncertainty visualizations were developed by Fisher (1993). He used the duration as a variable to represent uncertainty in soil classifications. Areas of low uncertainty were constantly shown in the colour of the respective soil type, while areas with high uncertainty dynamically changed colour. Fisher (1994) later even proposed the sonification of uncertainty information.

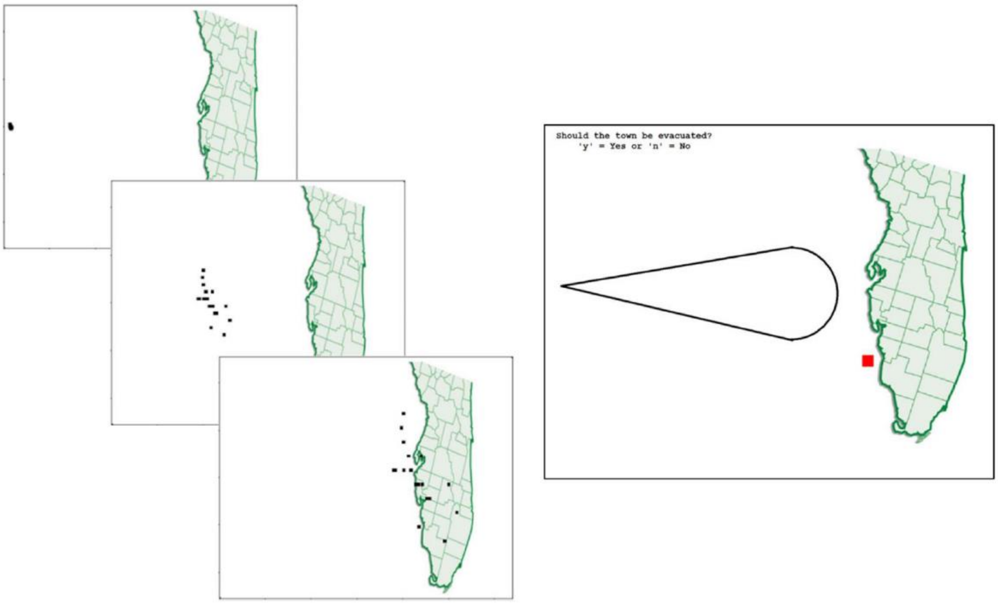


Figure 9: Dynamic uncertainty depiction for hurricane forecasts (left) compared to a simplified cone of uncertainty (right) as tested by Witt and Clegg (2021).

Witt and Clegg (2021) conducted a study to compare a static hurricane cone of uncertainty with a dynamic ensemble visualization method (see Figure 9). Through a series of experiments they found that participants were able to deal with the dynamic ensembles as they correctly extracted the information from the map displays. They also observed that some of the common misinterpretations tied to the hurricane cone of uncertainty – such as a boundary effect – can be avoided with their dynamic method. Therefore, they saw large potential for dynamic uncertainty visualization especially in the context of hurricane forecasts.

2.4.5. Uncertainty Visualization Studies in the Spotlight

The chapters above have briefly mentioned a variety of different uncertainty visualization results. Four studies, which served as a major source of inspiration for the study at hand, will be put into the spotlight in this next chapter.

2.4.5.1. Study Portrait: Cheong et al. (2016)

The study conducted by Cheong et al. (2016) investigated the influence of different uncertainty visualization methods in the context of a wildfire hazard scenario. They conducted a series of three experiments with varying levels of difficulty in terms of the decision-making task. A total of five uncertainty visualization methods (boundary, colour hue, colour value, transparency and texture) as well as a verbal uncertainty communication method were tested in the three experiments designed as a between-subject study. An example of each method can be seen in Figure 10. For each map, participants decided whether they would stay at their home (indicated by the cross in the map displays) considering the indicated burn likelihood at that location. The decision outcome, the response time as well as qualitative questionnaire answers including preference tasks were collected during the experiments. In the first experiment, the participants were given 30 seconds to make their decision. The second experiment added time pressure by only offering 5 seconds for participants to answer. The third experiment incorporated an additional task. This task included the memorization of sentences, which competed for the participants' attention. Additionally, participants were able to win money if they made correct decisions in all experiments. Due to its relevance to the study at hand, only results of experiment 1 will be discussed below.

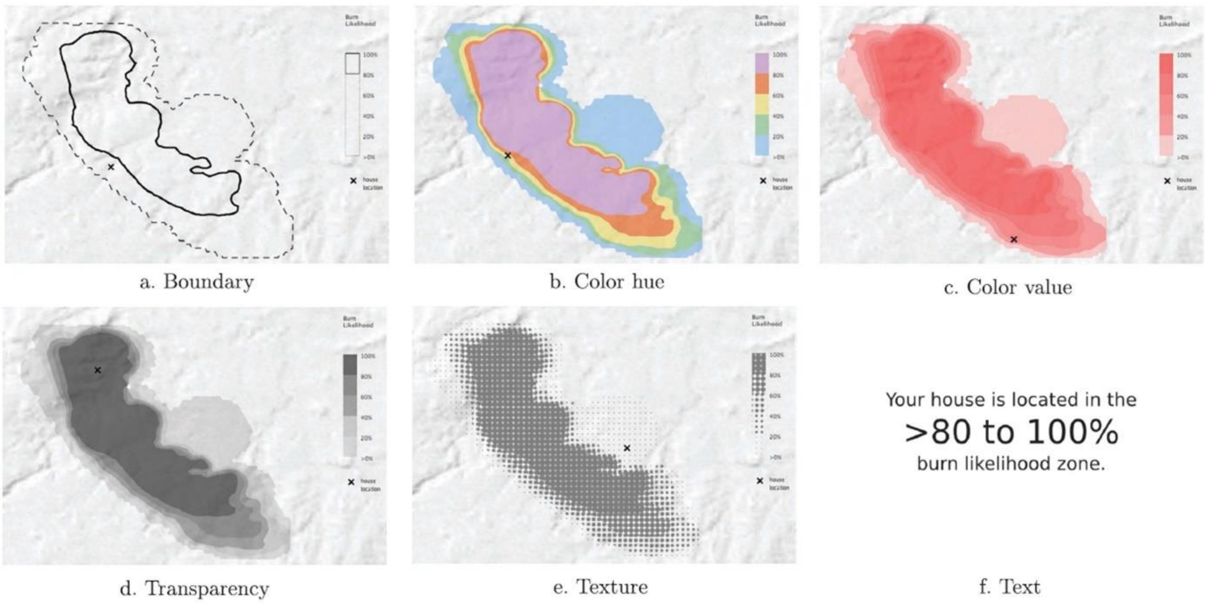


Figure 10: Uncertainty visualization methods tested by Cheong et al. (2016) (displays rearranged by the author).

A logistic regression analysis was performed to compare the results obtained with the different methods to the benchmark of the boundary method. It showed that participants were most likely to decide to stay at the house when uncertainty was visualized with colour hue, followed by texture, colour value, transparency, boundary and the text. However, these differences in the decisions were not significant. In terms of decision accuracy, the text method led to significantly more correct answers compared to the map-based methods. This method was followed by transparency, colour hue, colour value, texture and boundary in decreasing order. The response time was similar for all methods. One major difference was, however, found in terms of participants' preference for the methods. Colour hue ranked highest in this regard even though it was only on third place regarding the decision accuracy. Cheong et al. (2016) argued that this could be due to the high contrast between the hues, which facilitated the distinction of likelihood levels.

2.4.5.2. Study Portrait: Kübler et al. (2020)

Kübler et al. (2020) conducted a study on uncertainty depiction in hazard maps and posed a multicriteria decision-making task, which included the purchase of a house. By creating a within-subject design, they tested two different independent variables. Firstly, they wanted to investigate the effect of visualizing uncertainty in hazard map. So they created deterministic map displays which followed the official guidelines of Swiss hazard maps (more on these in Chapter 2.8). Other maps visualized the spatial uncertainty regarding the borders between the different danger zones. Secondly, when displaying the uncertainty, they tested the three uncertainty visualization methods of colour value, focus and texture (see Figure 11). The dependent variables recorded in the study were the decision outcome, response time as well as the participants' eye movement patterns. During the experiment the non-expert participants were asked to choose one of four houses displayed in the hazard maps according to their location, their price and the danger information. This implies that there was no distinctly correct answer to the task. The authors were much rather interested in a potential change in the decision outcomes with the different types of visualizations. Additionally, they put much focus on investigating the process of decision-making instead of only reporting on quantitative measures.

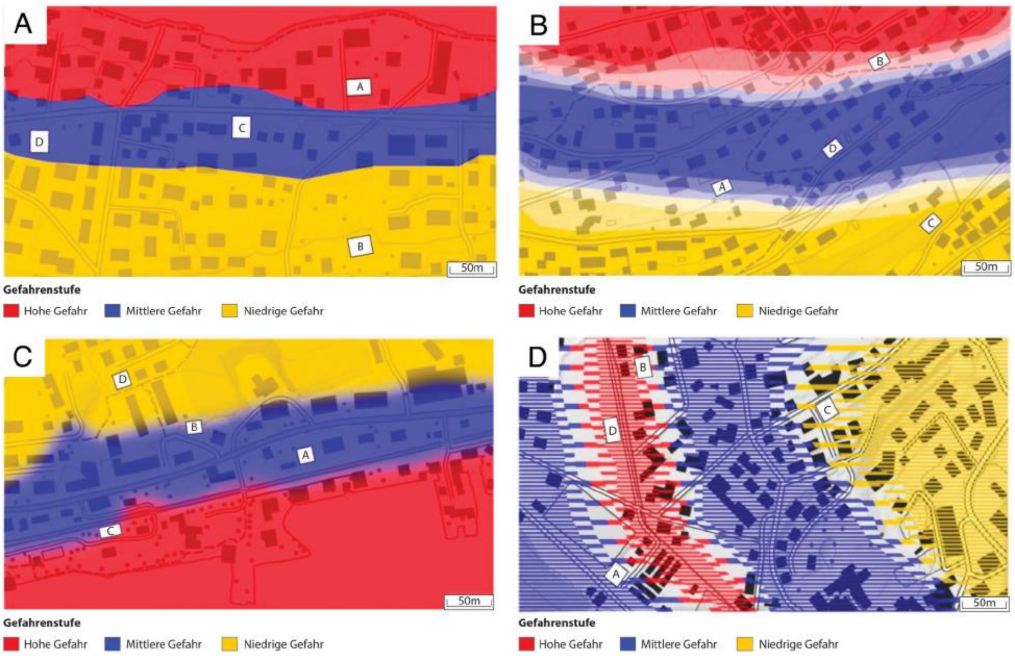


Figure 11: Uncertainty visualization methods tested by Kübler et al. (2020): A – no uncertainty, B – colour value, C – focus and D – texture.

They found that the inclusion of uncertainty information led to a change in the decision-making. Surprisingly, participants were more likely to decide on houses within the uncertain zones compared to areas where the danger zone was certain. Decision outcomes also differed depending on the type of uncertainty visualization. However, there was no consistent pattern observed and differences were not significant. Additionally, participants chose houses in the high danger zone more often when uncertainty was depicted. Houses at high danger locations were especially often chosen in the colour value visualization. The authors traced this observation back to a potential misinterpretation of colour value and suspect that participants interpreted the lighter values to represent a lower danger instead of a lower uncertainty regarding the choice of the danger zone. Therefore, it must be questioned whether the non-expert audience understood the uncertainty information correctly. Overall, the most important aspects in decision-making were found to be the characteristics of the house location no matter which visualization method was used.

2.4.5.3. Study Portrait: Ash et al. (2014)

The study by Ash et al. (2014) was conducted in the context of developing a new visualization for tornado warnings including uncertainty information. At the time of the study, tornado warnings were depicted deterministically through polygons, which extended over the area that was expected to be affected by the event within a short timeframe. However, considerable uncertainty is associated with the trajectory that a tornado takes in the future. They thus designed a between-subject study to test three different visualizations (see Figure 12). The first one represented the deterministic visualization, which was common practice at the time. The other two included uncertainty information through the visual variables colour hue – a spectral colour scheme – and colour value – a red gradient colour scheme. The authors made use of the semiotics associated with the hue in the red gradient design to convey the danger. The uncertainty information was communicated through different zones and the uncertainty increased with the distance from the current location. Participants were shown scenarios including a position and were asked to estimate their level of fear as well as how likely they would induce protective action if they were located at that position.

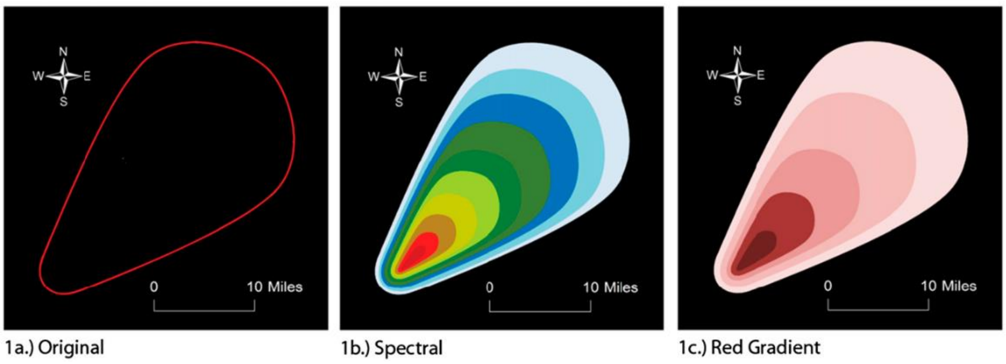


Figure 12: Uncertainty visualization methods tested by Ash et al. (2014).

They found that the answers for protective action exceeded the levels of fear for all visualization methods. The highest scores for both measures were recorded with the original design, the red gradient design took second place followed by the spectral design. Interestingly, the highest responses recorded with the original design were located in the centroid of the polygon. In the probabilistic designs, those were found in the lower left part of the visualization as one would expect. Therefore, the probabilistic designs enabled participants to correctly interpret the directionality of uncertainty, while no such aid was provided in the original design. Additionally, a boundary effect was observed. Participants were more likely to protect themselves if their position was located within the polygon shape. This effect was especially strong in the original design. It can be concluded from this study that the type of visualization has an influence on how people interpret the area under tornado threat and the spatial distribution of its uncertainty.

2.4.5.4. Study Portrait: Klockow-McClain et al. (2020)

Klockow-McClain et al. (2020) also investigated different ways of displaying uncertainty in tornado threat visualizations. The goal of the study was to find out whether the distance between a location and the tornado, the in-/exclusion of the location in the tornado polygon and the colour representing the uncertainty would have an influence on the decision outcomes. They tested deterministic designs of two different lengths as well as probabilistic ones, indicating the chance that a certain area would be affected by the tornado in percentages, in a between-subject design. Three colour schemes – sequential, divergent and spectral – were implemented for the probabilistic designs (see Figure 13). A control design included uncertainty information but no colour. Additionally, some participants were provided with verbal guidance during the task, while others were not. Participants were asked to indicate whether they would induce protective action at airport locations systematically placed on the maps.

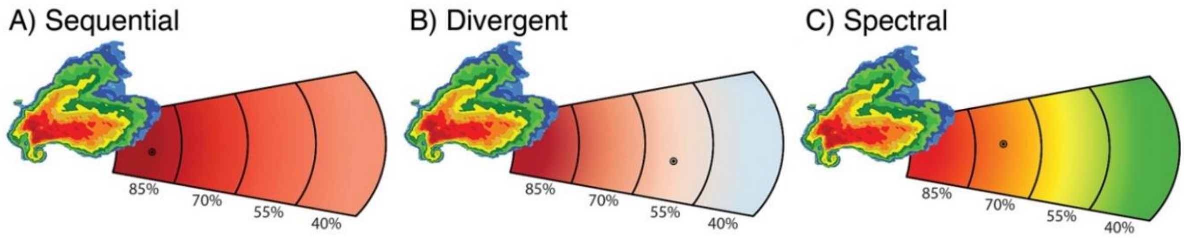


Figure 13: Uncertainty visualization methods tested by Klockow-McClain et al. (2020).

The results indicated that a boundary effect occurred again. Furthermore, a significant distance effect was observed for deterministic scenarios with fewer protective actions, the further away the location was from the storm. In the probabilistic scenarios, the responses were guided by the uncertainty information, showing that participants made use of this additional information. Overall, the probabilistic information led participants to protect more often if the probabilities were high and less often if they were low. This resulted in a higher decision accuracy compared to the deterministic scenarios. Sometimes, however, distance seemed to override uncertainty as the same values of numerical uncertainty led to different decision outcomes depending on the distance from the storm. The three colour schemes did not lead to significantly different responses. Yet, the sequential and the divergent colour scheme led to lower protective action if chances were low that a tornado would affect a location. The authors traced this result back to the saliency of the colour schemes which emphasize the areas of high chances. In terms of the verbal information it was found that trust in the forecast increased with the verbal guidance.

2.5. Decision-Making with Uncertainty

As it has become evident from the study portraits above, various studies in the realm of uncertainty visualization investigate the effect of uncertainty information on decision-making. Balleine (2007) described decision-making as the "ability of humans and other animals to choose between competing courses of action based on the relative value of their consequences". Map displays are very frequently used to support this process (Korporaal et al., 2020).

Various theories regarding the process behind decision-making have been established over the years. They can be broadly divided into two approaches: While some theories propose that humans are in fact capable of deciding rationally, others assume that decision-making occurs based on intuition with the help of so-called heuristics or cognitive shortcuts (Padilla et al., 2018). Studies on decision-making have clearly shown that human decisions frequently differ from the rational choice (Stanovich and West, 2000). Therefore, the consensus that humans make decisions both based on intuition as well as strategy has emerged in applied decision-making research. This is why the dual-process theory has gained in relevance (Padilla et al., 2018). The theory distinguishes between two types thinking: System 1 and System 2. System 1 is used for rather quick decision-making, often called *automatic thinking*. As these decisions are frequently made unconsciously, heuristics are applied to keep cognitive load at a low level (Ehrlinger et al., 2016; Stanovich and West, 2000). The use of heuristics is thus a precondition to decrease cognitive load of decisions and to enable quick decision-making (Joslyn and Savelli, 2021). System 2, on the other hand, is applied in more complex decision-making or *deliberate thinking*. It occurs consciously and requires larger cognitive effort compared to System 1 decisions. During System 2 decision-making, humans weigh different possible outcomes with the goal of making a rational choice (Ehrlinger et al., 2016; Joslyn and LeClerc, 2013). This does, however, not mean that System 2 results in completely rational decision-making. Errors can nevertheless occur due to the limited cognitive capacity (Joslyn and LeClerc, 2013). System 1 and 2 can also be seen as affective and analytic ways of thinking as described by Epstein (1994).

Figure 14 shows a model of how decision-making occurs with the help of a visualization. The model combines the dual-process theory with concepts of visualization processing (Padilla et al., 2018). It illustrates that visual processing and the subsequent decision include the matching of an object with previous knowledge stored in memory (MacEachren, 1995). The decision-making step at the very end of the process can be performed either with System 1 or System 2 processing. The difference between System 1 and 2 decisions is strongly tied to the amount of working memory involved in the decision-making process (Padilla et al., 2018). Specifically in decision-making with visualizations, System 1 is commonly applied during the first steps of information extraction, while System 2 is used when larger amounts of working memory are required to solve a decision-making task with the help of a previously acquired strategy (Padilla et al., 2018).

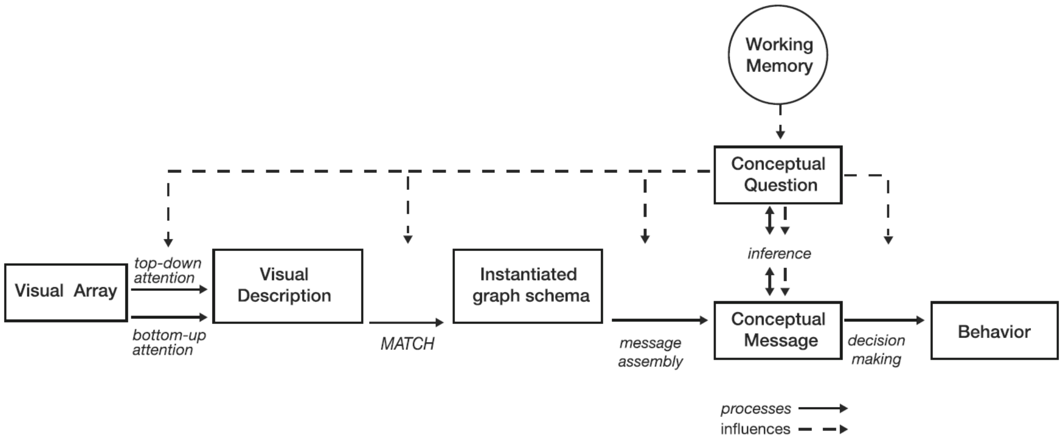


Figure 14: Model of decision-making with visualizations proposed by Padilla et al. (2018), which combines the dual-process theory of decision-making with a model of visualization processing.

The theory and the model presented above have shown that decision-making is a complex process. This complexity can increase even more if uncertainty is involved as it is the case in most decisions in real life (Hope and Hunter, 2007; Korporaal et al., 2020). Yet, knowledge on the quality of the information, upon which a decision is based, is an important precondition for informed decision-making (Gershon, 1998). The uncertainty information is included in decision-making by tying it to the likelihood of an option, which is then weighed against other possible outcomes (Joslyn and LeClerc, 2013; MacEachren et al., 2005). Thus, the addition of uncertainty into the decision-making process leads to an increase in the amount of information which needs to be processed (Kinkeldey et al., 2017). This is further challenged if uncertainty is communicated in complex formats (more on this in Chapter 2.6). Therefore, decision-makers need to simplify the problem when making decisions with uncertainty by focusing on a subset of the provided information and applying certain decision-making strategies (Korporaal et al., 2020). Research has shown that decision-makers are starting to accept uncertainty as an integral part of any decision-making task instead of trying to avoid or minimize uncertainty (Deitrick, 2012).

Various empirical studies have shown that humans deviate from rational probability theories, whose application is generally difficult, when judging uncertain events (Kahnemann and Tversky, 1982). To reduce the complexity of a decision in the face of uncertain information, a variety of heuristics can be applied (Tversky and Kahnemann, 1982). The use of heuristics can reduce the cognitive load of a decision with uncertainty (Joslyn and Savelli, 2021). It might come across as if these heuristics introduce a bias into decisions. Quite contrary, heuristics often support more accurate decision-making (Hullman et al., 2019). The three most important heuristics employed in decision-making under uncertainty are representativeness, availability and anchoring (Tversky and Kahnemann, 1974). For a more extensive overview of other heuristics see Ehrlinger et al. (2016) and Kahnemann et al. (1982).

The representativeness heuristic is framed as follows: Humans decide between uncertain events by judging whether an event X is representative of an event Y in X's parent population (Kahnemann and Tversky, 1982; Tversky and Kahnemann, 1974). Tversky and Kahnemann (1974) illustrate this heuristic with the following example: When asked to infer the occupation of a person who is described as shy, helpful and well structured from a selection of different options (including an airline pilot, a physician, a farmer, a librarian and a salesperson), one is likely to assume that this person is a librarian. It is the representativeness heuristic which leads one to guess that the mentioned characteristics are representative of a librarian.

The availability heuristic implies that it is easier to imagine the occurrence of an event if that type of event is more common in general. This heuristic leads humans to overestimate the likelihood of an event, if they can recall it easily (Hope and Hunter, 2007; Slovic et al., 1982). For instance, an individual would judge the likelihood of being affected by a natural hazard higher, if they had just been affected a few weeks ago. On the contrary, if they had not experienced a natural hazard in years, they could overestimate their feeling of safety.

The anchoring heuristic entails that estimating the likelihood of a certain event is influenced by the definition of an initial value – the anchor (Ehrlinger et al., 2016). This value is then used as a starting point to make relative decisions (Padilla et al., 2018). This heuristic leads people to "overestimate the probability of conjunctive events and to underestimate the probability of disjunctive events" (Tversky and Kahnemann, 1974). An example of this heuristic is provided by Ehrlinger et al. (2016): When asked to indicate the year George Washington became president, people use the year 1776 (declaration of independence of the United States) as an anchor to answer the question.

Lastly, a common strategy when dealing with uncertain information worth mentioning here is a risk aversion. It implies that humans would rather choose a certain option instead of taking a risk even if the gain of taking a riskier or more uncertain option is equal or higher (Joslyn and LeClerc, 2012).

Decision-making under uncertainty is often supported by geographic information in fields such as the insurance industry or hazard management (Kübler et al., 2020). Therefore, it does not come as a surprise that a rising number of studies are investigating how the visualization of uncertainty influences decision-making (Kinkeldey et al., 2017). As previously mentioned, the overall goal of uncertainty visualization is to enable informed decision-making. This chapter has shown that decision-making itself is a complex process, which can be further challenged if uncertainty is involved. The decision-making process with uncertainty visualizations is characterized by two processes. In a first step, the uncertainty information is extracted from a map display. Secondly, the extracted information is applied to infer on a specific task (Kusumastuti et al., 2022). Therefore, the important distinction must be made "between being able to identify information in a visualization and being able to use it effectively in a decision" (Hullman et al., 2019). Visualizations of uncertainty, aiming to support decision-makers, must thus be designed as simple as possible. Additionally, they must guide a map reader's attention to the information which is most relevant for the decision-making process (Pang, 2008).

2.6. Uncertainty Communication

Van Der Bles et al. (2019) argued that uncertainty can be communicated in one of three ways or a combination of them: visually, numerically and verbally. The visual communication of uncertainty was the subject of Chapter 2.4. the next chapter will focus on numerical and verbal forms of uncertainty expressions.

The fundamental goal of any form of communication is to convey information to people (Van Der Bles et al., 2019). This generally true, no matter if uncertainty is part of the communication or not. If uncertainty is communicated, a communicator is interested in the effects of this additional information on decision-making and behaviour. Through an expert survey Kunz et al. (2011b) found that natural hazard experts call for an improved communication of uncertainty information among experts. However, the experts also admitted that a consensus on how to go about uncertainty communication is still lacking. To support uncertainty communication, Van Der Bles et al. (2019) developed a framework to guide uncertainty communicators (see Figure 15).

In the first step of the framework, the communicators themselves are defined. These can usually take two roles: the role of a domain expert, who collects or creates data with uncertainty, or a communicator, such as a visualization expert. Different aspects of uncertainty such as its type and its sources, as discussed in Chapter 2.3, are relevant when defining the content of the communication. The forms of uncertainty communication include numerical or verbal expressions. A key aspect to define in this step is whether uncertainty is only communicated in the form of text or if it is also visualized. The next aspect is to consider the target audience of the uncertainty information (Van Der Bles et al., 2019). It has been shown that the information need of different actors in a risk assessment workflow – such as risk analysts, decision-makers and the affected public – varies significantly (Loucks et al., 2005). Lastly, expected effects of communicating uncertainty on the audience's cognition, emotion, trust as well as the resulting actions and decisions must be carefully considered (Van Der Bles et al., 2019).

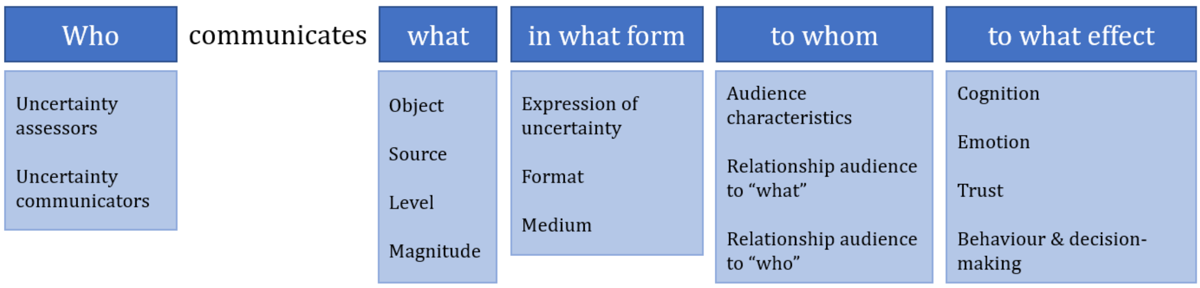


Figure 15: Deconstruction of the uncertainty communication process (figure designed by author, inspired by Van Der Bles et al. (2019)).

Likely one of the largest challenges of communicating uncertainty is thus taking into account all these aspects while preventing the information from becoming enormously complex. The risk of a cognitive overload due to the complex interpretation of uncertainty information must not be neglected (Doyle et al., 2014; Joslyn and LeClerc, 2013). One way to deal with the complexity of interpreting uncertainty lies in the development of a communication strategy which offers specific training for different actors (Doyle et al., 2014).

Choosing a suitable communication format is a major challenge. The appropriateness of a specific format is tied to the different steps of the communication framework presented above (Van Der Bles et al., 2019). The numerical uncertainty communication format traditionally takes the form of probability percentages or value ranges (Boukhelifa and Duke, 2009). The main advantage of numerical uncertainty expressions is their precision (Van Der Bles et al., 2019; Wallsten and Budescu, 1995). Bisantz et al. (2011) found that the inclusion of numerical uncertainty labels was a valuable addition to their visualizations. Yet, various authors have argued that numerical expressions in the form of percentages might not be the optimal form of communication. They propose that natural frequencies (e.g., 2 in 10 instead of a 20% chance) are much easier to interpret due to their correspondence with human cognitive processing of uncertainty (Hoffrage and Gigerenzer, 1998; Joslyn and LeClerc, 2013; Padilla et al., 2021a). Additionally, Jenkins et al. (2017) have argued that the numerical format could be interpreted as too precise and thus less credible, especially in the context of natural hazards.

Alternatively, uncertainty can be communicated through verbal expressions such as *unlikely* and *likely*. It can be argued that the vagueness of verbal expression supports the overall concept of uncertainty (Joslyn and Savelli, 2021). The verbal format was initially seen as suitable to communicate uncertainty to non-experts due to their lower complexity (Joslyn and Savelli, 2021). However, it has been found that the interpretation of verbal expressions is highly variable between subjects and is related to the personal vocabulary, which makes it difficult to standardize the terms (Budescu et al., 2012; Van Der Bles et al., 2019; Wallsten and Budescu, 1995). Furthermore, expressing distinct uncertainty with verbal expressions can be challenging due to their ambiguity (Erev and Cohen, 1990). Doyle et al. (2014) found that participants of their study tried to match vague verbal expressions to precise numerical ones before making a decision. This increased the overall cognitive load of the task.

Several text-based studies have investigated differences regarding the preference and the interpretation of numerical and verbal uncertainty expressions. Wallsten and Budescu (1995) reviewed a variety of studies and formulated a set of assumptions and principles in this context. When participants preferred verbal uncertainty, they did so because the expressions seemed natural and easy to interpret. The main reason to prefer numerical expressions was their precision. However, the review also showed that the quality of decision outcomes was not affected by the uncertainty format. Additionally, they noted that the context of a decision can have an influence on how an uncertainty expression is interpreted. A context effect for numerical and verbal expressions was also found by Windschitl and Weber (1999). A study on basketball game predictions conducted by Erev and Cohen (1990) aimed at gaining insight into which uncertainty communication formats are generally preferred by different actors. They found that most uncertainty communicators chose verbal expressions to convey the information. Decision-makers receiving this information, however, preferred to be presented with numerical uncertainty. This reinforced the existence of a so-called *communication mode preference paradox*. Olson and Budescu (1997) found that subjects strongly preferred numerical uncertainty expressions over verbal ones. If the uncertainty was communicated for certain events, the numerical format was preferred for both directions of communication. However, when events were vague, the preference for the numerical format was weaker. The communication mode preference paradox was only observed in 15% of the subjects for vague events. Therefore, they were able to show that although numerical uncertainty expression are usually preferred, none of the two formats wins it all when varying the uncertainty of events. Doyle et al. (2014) investigated the influence of the communication format (verbal vs. numerical) on evacuation decision in the context of volcanic eruptions. They found that that participants evacuated less often when presented with verbal statements, which were seen as more ambiguous. The difference in the decisions between the two formats was especially large for scientists compared to non-scientists. They traced this back to the high familiarity with numerical uncertainty in the scientific community.

One way to prevent a mismatch between how numerical and verbal uncertainty expressions are interpreted is to implement a translation table. This table clearly defines a numerical range of uncertainty which is represented by a specific verbal expression (Doyle et al., 2014). One example of such a translation table is the likelihood scale used by the Intergovernmental Panel on Climate Change (IPCC). The IPCC takes a key role in communicating scientific results on climate change – associated with vast uncertainties – to policy-makers (Kandlikar et al., 2005). To enable consistency in how uncertainty is communicated in their reports, the likelihood scale in Table 3 was developed. This scale can be applied to results stemming from modelling, quantitative analyses or expert opinions (Mastrandrea et al., 2010).

Table 3: IPCC likelihood scale for standardized uncertainty communication (Mastrandrea et al., 2010).

Likelihood	**Term**
99-100%	Virtually certain
90-100%	Very likely
66-100%	Likely
33-66%	About as likely as not
0-33%	Unlikely
0-10%	Very unlikely
0-1%	Exceptionally unlikely

The performance of the IPCC likelihood scale was investigated through different studies. Budescu et al. (2009) presented subjects with statements from an IPCC report containing verbal uncertainty expressions and asked them to assign a numerical estimate as well as a numerical uncertainty range to it. Some participants were provided with the translation table, while the control group was not. Some participants additionally saw numerical estimates next to the verbal terms. They found that answers of participants which had access to the translation table were more consistent with the likelihood scale compared to the control group. However, the overall consistency was still rather low. It was also evident that large between-subject differences in the interpretation of verbal terms existed, especially in the control group. By providing numerical estimates, the inconsistency between responses and the official scale decreased from 53% to 41%. Even though this decrease was observed, an inconsistency of 41% represents a considerable deviation between how participants interpret verbal uncertainty and how the likelihood scale intends them to do so.

In another study on the IPCC likelihood scale, Budescu et al. (2012) aimed at applying a similar concept on a larger sample, which was representative of the general public. Again, participants were divided into the control, translation and verbal-numerical groups. The task was the same as above.

Table 4: Results of Budescu et al. (2012) showing the mean numerical uncertainty estimate [%] per IPCC term compared to the likelihood range [%] of the IPCC likelihood scale.

	IPCC Term			
	Very unlikely	**Unlikely**	**Likely**	**Very likely**
IPCC likelihood range [%]	0-10	0-33	66-100	90-100
Mean numerical estimate [%]	41	44	54	62

The results of how participant matched numerical likelihood estimate to the verbal expressions in Table 4 show that large inconsistencies were observed once more. The overall consistency for the verbal-numerical group were significantly higher compared to the other two groups. Again, large differences between individuals were recorded. Additionally, they found that subjects with a stronger belief in global climate change assigned significantly higher numerical estimates to three out of four verbal expressions. They argued that providing participants with a dual scale, composed of verbal and numerical uncertainty expressions, would perform best. However, it remains open to debate whether the combined communication format is actually beneficial. A recent review of uncertainty communication studies by Dhami and Mandel (2022) has touched on first results indicating that this might not be the case.

2.7. Including or Excluding Uncertainty Information

All the topics discussed so far culminate into the question of whether uncertainty can be successfully visualized and communicated when reporting scientific results or not. Some key arguments for the inclusion as well as the exclusion of uncertainty information are presented below.

2.7.1. Arguments to Include Uncertainty Information

Generally, the inclusion of uncertainty information can make decision-making more transparent and enable informed decision (Thompson and Warmink, 2016). This has the potential to increase decision quality. Despite initial doubts, public trust in scientific information can actually be enhanced due to the transparency resulting from communicating uncertainty (Van Der Bles et al., 2019). Additionally, Hullman (2020) found that scientists feel responsible to include uncertainty information. Furthermore, it is argued that communicating uncertainty should be part of a sound scientific workflow, as it allows more comparability between results and supports scientific transparency (Pappenberger and Beven, 2006). Joslyn and LeClerc (2012) argue that even non-experts are aware that uncertainty is an inherent part of the world. Therefore, it is assumed that humans understand expressions of uncertainty in practice, even if theoretical knowledge on the concept of uncertainty might be lacking (Joslyn and Savelli, 2021).

Joslyn and LeClerc (2012) and Nadav-Greenberg and Joslyn (2009) both investigated the effect of communicating uncertainty in a road salting scenario. Both studies found that the inclusion of uncertainty in temperature forecasts led to higher decision accuracy. Additionally, the trust put into the forecast increased when participants were provided with uncertainty information (Joslyn and LeClerc, 2012). Nadav-Greenberg and Joslyn (2009) included the opportunity to request uncertainty information in one of their experimental groups. This option was frequently made use of. These results showed that non-experts acknowledge the value of uncertainty information in a decision-making task.

MacEachren et al. (1998) was among the first to report on positive results regarding the inclusion of uncertainty information in map displays. They presented health statistics and their uncertainty coincidentally using the visual variable of colour with a texture overlay. The results showed that participants were able to cope with this additional information. Similarly, positive results for depicting uncertainty in map displays were recorded for instance by Bisantz et al. (2011) and Leitner and Buttenfield (2000). These results showed that the additional information of uncertainty does not lead to an increased complexity of map displays. Kinkeldey et al. (2017) conducted a review of 87 studies investigating the effects of visualizing uncertainty. The majority of these studies found that the visualization of uncertainty influenced decision accuracy positively. Additionally, it was found that decision time did not significantly increase, despite the higher cognitive load. Overall, their review showed that the graphical depiction of uncertainty outperforms solely representing uncertainty through numbers or text.

2.7.2. Arguments to Exclude Uncertainty Information

Hullman (2020) investigated the reasons why authors exclude uncertainty in visualizations by surveying 90 visualization authors and conducting interviews with visualization experts. Around 30% of the surveyed authors included uncertainty in 10% or less of their visualizations. 62% of the authors stated that a reason to exclude uncertainty is to prevent the viewers from being overwhelmed. The notion that non-experts cannot understand uncertainty is extremely prevalent in literature. Therefore, experts argue that decision-makers should be provided with certain information as their lack of numeracy leads to false interpretations (Fischhoff, 2012). Similar concerns have been expressed specifically for the visualization of modelling results

(Pappenberger and Beven, 2006) and the domain of natural hazard predictions (Kunz et al., 2011b). A further argument to exclude uncertainty is a fear of communicating it from the expert's perspective. If it is assumed that decision-makers want to work with the 'truth'. Thus, experts fear that including uncertainty would make them appear incompetent and reduce the trust between them and the decision-makers (Fischhoff, 2012; Van Der Bles et al., 2019). The last major argument why experts tend not to include uncertainty information is a lack of knowledge on how to do so. The synthesis by Fischhoff (2012) as well as Hullman's (2020) survey have shown that experts struggle to find fitting formats on how to represent uncertainty information verbally, numerically as well as visually.

Likely one of the first empirical studies in the realm of uncertainty visualization was conducted by Schweizer and Goodchild (1992). They chose a coincident visualization method and combined colour saturation and colour value to represent both the data and its uncertainty. They concluded that participants were not able to differentiate between the two visual variables. Additionally, they found a rather low decision accuracy in absence of a legend which they traced back to the fact that map readers were not used to uncertainty depictions. Nevertheless, they argued that participants understood the uncertainty information if a legend was present. In another study already mentioned above, Hope and Hunter (2007) found that the inclusion of uncertainty information in the form of glyphs led to irrational decision-making. As their participants preferred low uncertainty zones over high uncertainty zones even if the latter was more suitable in their airport siting scenario, they argued that humans prefer to make decisions based on certain data. They thus concluded that humans cannot logically process uncertainty information.

The arguments presented in the past two subchapters clearly show that a discrepancy between the concepts of uncertainty visualization and the visualization practices persists. Even though it is generally agreed that uncertainty is an inevitable part of any type of data, it is often not communicated or visualized for the public (Joslyn et al., 2007; Joslyn and Savelli, 2021). However, as shown by Kinkeldey et al. (2017), the majority of research has proven that decision-makers can deal with uncertainty information in map depictions. Therefore, the study at hand was developed under the assumptions that the depiction of uncertainty has positive effects on map-based decision-making.

2.8. Natural Hazard Mapping

The effectiveness of uncertainty visualizations can be highly application-specific (Boukhelifa and Duke, 2009). The field of application relevant for the study at hand are map displays in natural hazard management. Three terms must be defined in order to discuss natural hazards in a well-rounded way: hazard, vulnerability and risk. A hazard describes a natural phenomenon including its potential damages in a given temporal and spatial context (Graf et al., 2019; Raetzo et al., 2002). The general characteristics of a hazard are its probability, intensity and location (Zimmermann et al., 2005). Vulnerability concerns the potential consequences of a natural hazard event including the exposure of property or infrastructure (Chesneau et al., 2005; Zimmermann et al., 2005). Risk is then often interpreted as the multiplication of hazard and vulnerability representing the possibility of an event including the potential negative consequences (Ramsey, 2009; Zimmermann et al., 2005).

Natural hazard management, which includes the conduction of hazard assessments, describes a method to deal with the potential occurrence of natural hazards. Its goal is to prevent damages by identifying trends regarding future natural hazard events (Kunz et al., 2011a, 2011b). Parameters which are assessed through natural hazard modelling include the location of an event as well as the magnitude, the duration and the time of the event (Thompson and Warmink, 2016).

Hazard maps are a valuable tool to visually communicate the results of hazard assessments. Their development has become a standard practice in a variety of countries including Switzerland (Kunz et al., 2011a). Kunz et al. (2011a) have noted that the expectations posed towards hazard maps have increased lately. While the hazard maps used to be informed by past events documented in registers, the implementation of natural hazard models has become an important part of the hazard map production today (Walser et al., 2014; Zimmermann et al., 2005). Due to the variety of stakeholders involved in natural hazard management, it is very important that hazard maps are easy to understand (Chesneau et al., 2005), since it has been found that people only react to hazards if they actually perceive them as such (Slovic et al., 1982).

Different types of hazard maps such hazard indication maps (including the hazard type and its spatial extent) and danger maps (additionally showing the intensity and probability of an event) have been developed in the Swiss framework of natural hazard management. The first Swiss hazard maps range back to the 1970s and they are generally created at scales between 1:5'000 and 1:10'000 (Zimmermann et al., 2005). The overall goal of hazard maps is to define a consensus on hazard assessments across various physical processes (ARE and FOEN, 2005). The Swiss Cantons are legally required to create hazard maps for a range of natural hazards including avalanches, flooding and gravitational processes (FOEN, 2015). Based on a combination of intensity and probability categories, three danger levels were defined. They are visualized in red, blue and yellow hues (see Figure 16). A forth level, indicated through yellow stripes, represents a residual danger (FOEN, 2015; Raetzo et al., 2002). These hazard maps support decision-making in application fields such as spatial planning, emergency planning and the development of protection measures (Zimmermann et al., 2005).

An analysis of the Swiss hazard mapping framework and its application in other countries by Zimmermann et al. (2005) showed that the system is perceived as transparent and easily understandable. However, they also found that the system can be very complex, which is why decision-makers should be more involved in the creation of hazard maps. Regarding the standard colour scheme, they noted that the blue danger level was often misinterpreted as flooding and not perceived as the medium danger level (Zimmermann et al., 2005). Weaknesses of the colour scheme were also mentioned by Kübler et al. (2020) and Kunz et al. (2011a). These were mainly tied to the lack of an intuitive order of the used colour hues (Kübler et al., 2020).

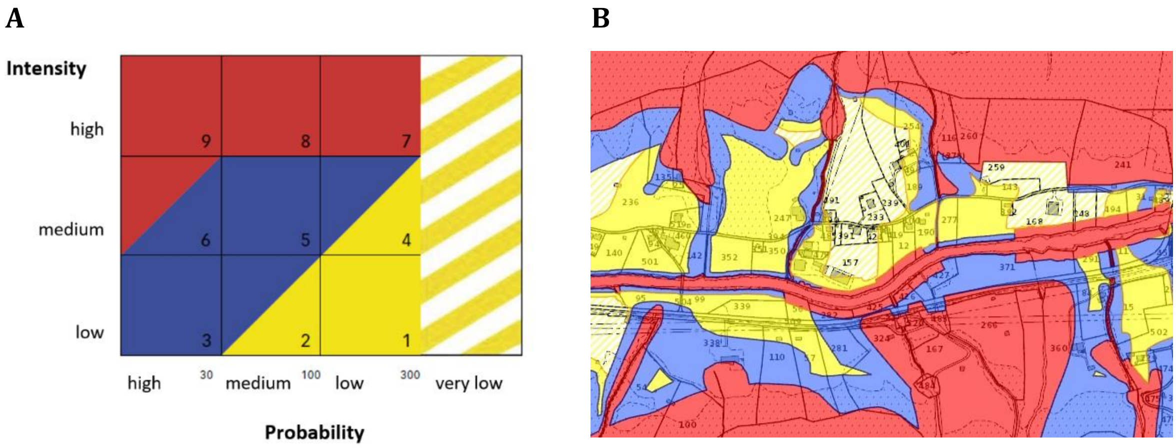


Figure 16: A: Danger levels in the Swiss hazard maps as a combination of intensity and probability of a potential event (ARE and FOEN (2005), labels translated by the author). B: Example of a hazard map for the municipality of Lütschental BE (Geoportal des Kantons Bern, 2019).

One major drawback of the Swiss hazard mapping framework, and quite frankly for the field of natural hazard management as a whole, is its lack of uncertainty information (Kubicek et al., 2012; Kübler et al., 2020). Up until now, hazard maps display the borders between danger zones as solid lines even though it is generally agreed that these borders are subject to major uncertainty (Trau and Hurni, 2007). This is tied to the inherent uncertainty in natural hazard assessments regarding the intensities and probabilities of events (FOEN, 2016). The key problem preventing the inclusion of natural hazard uncertainty can be found in the missing cartographic guidelines to visualize it (Kunz et al., 2011a). The study by Kübler et al. (2020), discussed in detail above, was a first step towards empirically investigating how uncertainty could be incorporated into the current hazard mapping framework.

2.9. Debris Flows

The Swiss hazard maps are collectively created for a variety of natural hazards (ARE and FOEN, 2005). The natural hazard focused on in this study is the debris flow, a mass movement which frequently occurs in the Swiss alps. Debris flows are a natural process during which large volumes of water-entrained sediment travel along established channels (Hirschberg et al., 2021a) (see Figure 17). They generally have a rather high density as the sediment content can be between 30% and 70% (BWW et al., 1997). The speed of debris flows can reach up to 10m/s (Raetzo et al., 2002). The flows are triggered by large volumes of water originating from heavy rainfall or glacial lake outbursts for instance (Frank et al., 2019; Stoffel et al., 2014). One important precondition for a debris flow to occur is the availability of sediment. Therefore, the volume of a potential debris flow is strongly dictated by the amount of sediment found in the initiation zone. Sources of this sediment include rock avalanches, landslides as well as rock glaciers, which are partially collapsing (Frank et al., 2019). As mentioned above, debris flows follow existing channels. However, when entering flatter terrain, they frequently reach regions outside of the steepest area and deposit sediment beyond the channel (Huggel et al., 2003). Another aspect which can lead to sediment deposition outside the channel is the erosion of the riverbed itself (Graf et al., 2019). The volume of a debris flow is thus composed of the initial volume as well as additional volume accumulated over the course of an event (Frank et al., 2019). The large volumes and speeds of debris flows result in the high damage potential of this process. On the one hand, its erosive forces can destabilize the channel. On the other hand, the forces at the front of a debris flow, where large rocks are transported, can cause major damages to infrastructure, buildings and humans (BWW et al., 1997).

Figure 17: Approaching debris flow at the WSL debris flow observation station at Illgraben (VS) on the 28th of July 2006 (Image: Corina Gwerder (WSL, ETH)).

Due to the large damage potential of debris flows, it is of high importance to identify areas which could be affected by future events (Schraml et al., 2015). Therefore, efforts to predict future debris flow events have been made in the field of physical geography. The most important parameters necessary to accurately predict a future event are the magnitude, the frequency and the timing (Frank et al., 2019). Records of past events serve as a source of information to estimate these parameters. However, parameters such as the spatial extent and the velocity are often missing for past events. At the majority of debris flow site, only extreme events were systematically analysed (Graf et al., 2019). Determining the composition of the different materials, which has a large influence on the behaviour of a debris flow, is another aspect which makes their prediction challenging (Graf et al., 2019).

Another method to predict future debris flow events and their respective spatial extent is the application of simulation models (Schraml et al., 2015; Walser et al., 2014). The models have likely become the most popular way of predicting natural processes, even though they contain major uncertainty themselves (Liu et al., 2017). The use of these models requires knowledge on the model parameters. For other natural hazards, such as snow avalanches, substantial knowledge regarding those parameters exists. This is unfortunately not the case for debris flows (Schraml et al., 2015). One key parameter, which must be defined in debris flow modelling, is the volume, whose estimation is tied to large uncertainties (Walser et al., 2014). Additional challenges arise from insufficient spatial and temporal resolutions of digital elevation models (DEMs), which do not accurately represent the constant changes in debris flow channels (Graf et al., 2019).

A further aspect which challenges debris flow prediction and introduces additional uncertainty is the ongoing climate change. Research on climate scenarios has shown that, while precipitation is expected to decrease in summer, more precipitation can be expected during the remaining seasons of the year (Hirschberg et al., 2021b). Additionally, precipitation events are expected to become more extreme (IPCC, 2013; Stoffel and Huggel, 2012), which could have an effect on debris flow triggering. The rising temperatures which lead to the thawing of permafrost are especially relevant in alpine regions (IPCC, 2013). This is expected to increase sediment availability for debris flows in permafrost regions (Hirschberg et al., 2021b). The consequences of climate change could lead to unprecedented debris flow events (Graf et al., 2019). The changes in temperature as well as precipitation could have an influence on debris flow volumes and return periods. In their study investigating climate change induced alterations in the debris flow process for a site in the Swiss valley of Mattertal, Stoffel et al. (2014) found that the debris flow season could be extended in the future. Additionally, they expect that the event volumes could increase during the months of summer and autumn due to higher sediment availability from active permafrost layers. However, since a decrease in the precipitation is expected during summer, debris flow events could become less frequent during this season.

Since debris flow prediction results are frequently used as a basis for decision-making in the form of hazard maps, their uncertainties need to be discussed (Graf et al., 2019). Furthermore, the uncertainty inherent to the process of debris flows and the enhanced uncertainties due to climate change must be included in cartographic debris flow depictions. The only known study which implemented uncertainty visualizations on debris flow assessments is by Kunz et al. (2011a). They implemented an interactive web-based system which allows a broad range of stakeholders to investigate natural hazard assessments for different processes such as debris flows and snow avalanches. The functionality of the system and the cartographic presentation of the information were informed by experts in the field of (web)cartography. Besides the natural hazard assessment data, users can include uncertainty information in various different ways. These include numerical uncertainty values in a tooltip, displaying uncertainty in a separate map or including uncertainty in a bivariate way. Uncertainty can be visualized through colour value or bar glyphs in the univariate displays. The following uncertainty visualization methods were offered in the bivariate displays: density of an overlaid dot texture, contour lines, saturation or transparency (for an example see Figure 18).

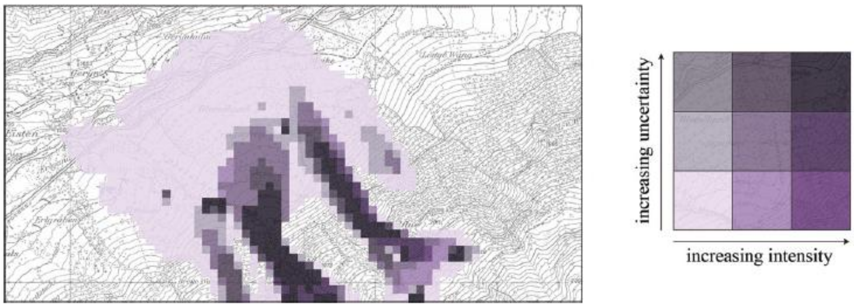


Figure 18: Bivariate uncertainty depiction for a snow avalanche event using colour value to represent the intensity and colour saturation for uncertainty (Kunz et al., 2011a).

The authors conducted interviews with different natural hazard experts to evaluate their system. These interviews showed that experts appreciated the functionality such as being able to change the visualization methods and colour schemes. Additionally, they reported that experts valued the possibility to switch the uncertainty information on and off. Experts were confident that they were able to properly interpret the additional information. When the goal is to draw attention to regions of high certainty, transparency was the preferred uncertainty visualization method, whereas when the focus needs to be on uncertain areas, saturation was seen as most suitable. Overall, experts saw it as an advantage that different uncertainty visualization methods were offered. The lack of an empirical evaluation of the system is, however, seen as a drawback of the study.

3. Methods

3.1. Expert Interviews

During the preparation of this thesis six interviews with experts working in the field of debris flow research and natural hazard management were conducted. The purpose of these interviews was to gain insights into these fields and acquire knowledge based on the experts' experiences. Additionally, current practices in developing hazard maps as well as debris flow modelling and mapping were investigated. Table 5 shows an overview of the demographic information of the interviewed experts. Their age ranged from 31 to 56 years (mean: 45.83, standard deviation (SD): 4.56) and they had an average of 16.10 years (SD: 6.94) of experience in their respective fields. The anonymized transcripts of the interviews can be found in the Appendix A to F. This chapter includes a short synthesis of the interviews, which served as a motivation and inspiration for the methods applied in this study.

Table 5: Overview of demographic information regarding the interviewed experts.

ID	Gender	Age [years]	Experience [years]	Background
E1	Male	56	19	PhD in Geology
E2	Male	31	3.5	MSc in Environmental Engineering
E3	Male	49	20	PhD in Geography
E4	Male	44	10	PhD in Geography
E5	Male	48	22	Forest Engineering, Construction Management
E6	Female	47	22	MSc in Geography

The three scientific experts agreed that debris flow events are currently not systematically mapped. However, they considered debris flows to be covered quite well by the Swiss hazard maps, which are frequently updated after larger events (see E1, E2, E3 in Appendix A, B, C). Nevertheless, E1 saw cartographic visualizations as a key instrument especially for communicating scientific findings to a non-expert audience.

Regarding the uncertainties associated with debris flows, experts concurred that the spatial uncertainty regarding the extent of a debris flow is the most important. They agreed that it is a challenge to estimate the volume of a potential debris flow (see E1, E2 and E6 in Appendix A, B, F). Expert E1 mentioned that "where they will flow is somewhat straightforward because they follow topography quite strongly. But how big they will be is very difficult to estimate in advance." The spatial extent of a debris flow is highly influenced by the volume as larger flows can leave the debris flow channel and reach unexpected areas. Another aspect of debris flows associated with uncertainties are the parameters of prediction scenarios such as the exact catchment and precipitation patterns (E3).

Even though these uncertainties are present in debris flow prediction, they are currently not visualized in hazard maps (E1, E2, E4, E5 and E6 see Appendix A, B, D, E, F). When contemplating whether uncertainty should be visualized in hazard maps or debris flow predictions, the practitioners emphasized that the simplicity of the maps was important. E4 mentioned that the purpose of hazard maps is to create awareness of existing hazards. E5 brought up that the simple visualization of current hazard maps enables quick decision-making. Nevertheless, they mentioned that judging situations at the border of the hazard levels can be rather challenging. They acknowledge that "the truth is continuous and the map shows a categorised order. One has to make a cut somewhere, which is not very easy" (E6, translation by the author). Additionally, it was stated that the standards in creating hazard maps have significantly increased during the past decades. This is related to the technological development in natural hazard modelling (E6). The same observation was also emphasized by Kunz et al. (2011a).

As uncertainty is not visualized at the moment, it was investigated whether it was considered during the process of creating hazard maps. On the one hand, experts mentioned that uncertainty in debris flows is frequently a topic of discussion during the hazard map creation or consultation meetings (E3, E4, E5 and E6). E3 stated that "it is almost always a topic of discussion: the uncertainty associated with certain processes and how one can deal with it" (translation by the author). On the other hand, E5 argued that decision makers also simply trust that hazard maps were created diligently and that uncertainty is sometimes not taken into account closely enough. They also mentioned that uncertainty can be misinterpreted by map users. For instance, E5 stated that non-experts expect a 100- or 300-year scenario not to occur during the next 100 or 300 years. This shows, that the understanding of probabilities might be incomplete.

Overall, the members of the scientific community, agreed that visualizing uncertainty in debris flow depiction would be of value (E1, E2 and E3). Some practitioners also showed interest in uncertainty visualization and assumed that people would understand it (E4 and E5). However, the experts also emphasized that there could be potential challenges. E3 mentioned that they support the inclusion of uncertainty in debris flow depictions but that it could be challenging to implement this in the spatial planning processes in Switzerland. Since public decision makers rely on the clear hazard categories, they suggest that uncertainty should only be considered as an additional source of information. E1 added that it would certainly need some training for practitioners and decision makers if uncertainty was depicted, but that it would ultimately lead to more objective decision outcomes.

Nevertheless, some reasons to exclude uncertainty in debris flow depictions and hazard maps were mentioned. E4 stated that the process of creating hazard maps in their current form is already rather complicated. E6 added that it would increase the complexity of this process even further. They also argued that it is the overall order of magnitude which is relevant as one can never predict a natural process in its entirety. Similarly, E5 brought up that more detailed calculations and models including uncertainty would narrow the personal responsibility of hazard map users. Additional ambiguities could emerge in the context of insurances and the granting of mortgages (E1). E4 summarizes that "the sharpness is necessary [in hazard maps], even through it is not completely true" (translation by the author).

In conclusion, the expert interviews allowed the author to dive into the topic of the thesis and view it through the eyes of experienced experts. There seems to be a lot of interest in including uncertainty in natural hazard maps. However, practitioners seem to be more reluctant and less optimistic than members of the scientific community in this regard. Overall, the relevance of uncertainty in natural hazards and cartographic depictions of these processes are a topic on the forefront of discussions in science as well as in practice.

3.2. Study Design

After the expert interviews, the concept of the study at hand was developed. To answer the research questions posed in Chapter 1.3, an empirical experiment was designed. The study was structured into three different parts (see Figure 19). In a first section, participants were asked to fill out a pre-test questionnaire to gain information on their demographics and previous knowledge. Additionally, this part included a thematic introduction into the topic of debris flows to guarantee that all participants had a basic understanding of the process.

During the main experiment, map displays were tested on the participants. The experiment was set up as a 2x2 mixed design. One advantage of mixed designs is that potential interactions between the levels of independent variables can be investigated (Martin, 2008). Therefore, two independent variables were actively manipulated in the experiment. The within-group variable is represented by the uncertainty visualization method (single-hue vs. multi-hue colour scheme). The method for uncertainty communication through the legend (numerical vs. verbal expressions) is the between-group variable of the study. Therefore, participants were presented with both types of uncertainty visualization methods but only one type of uncertainty communication method depending on which communication group they were assigned to (Martin, 2008). The participants were then asked to solve decision-making tasks, which involved the estimation of the potential damage caused by an uncertain future debris flow event. The environment of the experiment is considered to be a control variable, since the circumstances were kept constant for all participants in terms of the location, the hardware, the lighting conditions, the procedure and the conductor of the experiment.

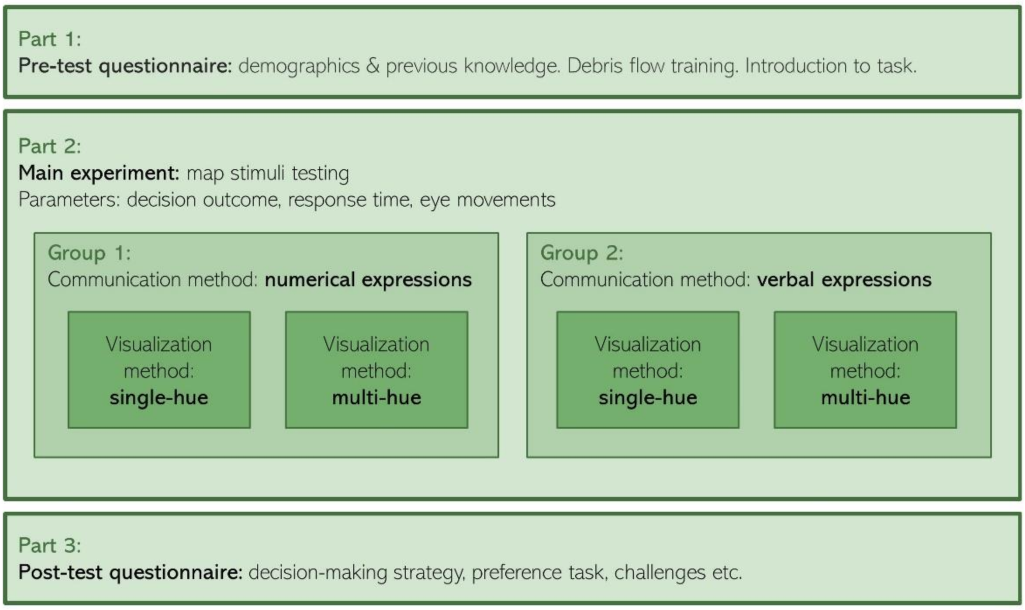


Figure 19: Overview of the study design and the experiment procedure.

The study was rounded off with a post-test questionnaire at the end of the experiment. The goal of this questionnaire was to collect qualitative data on the decision-making strategy employed by participants and other aspects through self-reports. This last part was considered especially important as uncertainty visualization studies have traditionally put a lot of weight on parameters such as decision accuracy without trying to understand why certain effects were observed (Hullman et al., 2019). As it happens, the collection of qualitative information through self-reports have been found extremely valuable to some more recent studies (e.g., Korporaal et al., 2020 and Kübler et al., 2020). After their review of studies investigating uncertainty depiction and its effect of decision-making, Kinkeldey et al. (2017) endorse the type of empirical approach taken here, which combines quantitative and qualitative methods.

3.3. Pre-Test

The pre-test questionnaire was twofold. On the one hand, its goal was to collect demographic information on the participants as well as their previous knowledge and map-reading skills. On the other hand, the pre-test questionnaire included a short training session on the process of debris flows, the introduction of the study scenario and the task which was later solved during the main experiment. The full pre-test questionnaire can be found in Appendix M. It was implemented with the survey tool LimeSurvey (LimeSurvey GmbH, 2021).

Demographic parameters included the age, gender and the maximum education level. Participants were also asked to indicate whether they worked in a field related to the study (such as natural hazard management or geographic information science/systems (GIS)). Around 10% of males and 1% of females suffer from some type of deficiency in their colour vision (Ware, 2013). Thus, participants were asked about their vision. To assess the so-called scientific literacy, which describes "an understanding of concepts such as a model and probabilities" (McMahon et al., 2015), participants were asked to judge their experience with topics related to the study such as map reading, data uncertainty and GIS. Lastly, Miran et al. (2019) suggested that having experienced a specific natural hazard could have an effect on how participants judged the threat of the hazard. Therefore, participants were asked to indicate if their had ever been personally affected by a debris flow.

The training session on the process of debris flows (see Appendix N) was included as Tak et al. (2014) argued that it is important to ensure that participants share a common knowledge base when working with a non-expert audience, even if the scenario tested in the study is simple. It was expected that participants had a basic understanding of the properties of a debris flow as well as its triggers and the uncertainty tied to it after this introduction. Lastly, the study scenario was presented to the participants (see Appendix O). They were introduced to the key components of the experiment such as the decision-making task, the damage scale and saw an example of a map stimuli to get familiar.

Frequently, studies make use of standardized tests to asses spatial ability or stress level of participants (e.g., Korporaal et al., 2020; Kübler et al., 2020). However, as these topics do not specifically relate to the research questions of this study, these types of tests were not used.

3.4. Main Experiment

The goal of the main experiment was to gain knowledge on potential changes in decision-making due to the way uncertainty was visualized and communicated. Therefore, participants were asked to solve a decision-making task with the help of map displays showing debris flow modelling results including the spatial uncertainty. Each participant solved the same task for 40 map stimuli. The measures which were recorded during this part of the study were the decision outcome through damage estimates and the response time. These measures represent the dependent variables of the study (Martin, 2008) and are classic measurements in uncertainty visualization research (Kinkeldey et al., 2017). Response time can be used as a proxy for how thoroughly a task is performed or how difficult a task is (Martin, 2008). Additionally, eye tracking technology was implemented to measure eye movements. This is a popular method to objectively gain insight into cognitive processes of map reading (Brus et al., 2012; Salvucci and Goldberg, 2000).

3.4.1. Development of Map Stimuli

Definitely the most important part of the main experiment and the study as a whole was the creation of the map displays of debris flow predictions with the associated spatial uncertainty. The essential steps involved in the development of the map stimuli are explained in the next few subchapters.

3.4.1.1. Debris Flow Uncertainty Data

Debris Flow Locations

Eleven Swiss debris flow channels were chosen to conduct process modelling for the map stimuli. The locations were chosen based on existing debris flow studies or via a recommendation by the interviewed debris flow experts. Table 6 gives an overview of the locations. Location LTB-T (Lauterbrunnen 2 (Rybibach)) was used as an example in the task introduction as well as two trial tasks before the main experiment started. This location was not tested during the main experiment. The other ten locations were used in the map stimuli of the main experiment.

Table 6: Debris flow locations used for modelling, their coordinates and the modelled debris flow volume range [m³].

Location (Channel)	**ID**	**Coordinates (LV95)**	**Volume Range [m³]**
Blenio (Riascio)	BLE	2'715'781, 1'159'033	5'000-50'000 (Frank et al., 2019)
Bondo (Bondasca)	BON	2'763'175, 1'133'470	20'000-200'000 (Frank et al., 2019)
Brienz (Glyssibach)	BRZ	2'646'457, 1'179'373	10'000-100'000 (Scheidl et al., 2008)
Guttannen (Rotlauwi)	GUT	2'666'302, 1'167'086	15'000-150'000 (Stoffel and Huggel, 2012)
Guttannen 2 (Spreitlauwi)	GUT2	2'663'605, 1'167'164	10'000-100'000 (Frank et al., 2019)
Lauterbrunnen (Gryfenbach)	LTB	2'635'360, 1'160'748	2'000-20'000 (Hitz et al., 2014)
Lauterbrunnen 2 (Rybibach)	LTB-T	2'636'420, 1'160'765	5'000-35'000 (Oberingenieurkreis I, 2014)
Leuk (Illgraben)	LUK	2'614'884, 1'126'818	5'000-50'000 (Hirschberg et al., 2021a)
Lütschental (Glattbach)	LUT	2'638'471, 1'165'597	1'000-10'000 (Scheidl et al., 2008)
Randa (Dorfbach)	RND	2'627'284, 1'105'929	1'000-10'000 (Deubelbeiss and Graf, 2013)
Silenen (Schipfenbach)	SIL	2'694'692, 1'182'429	5'000-50'000 (Frank et al., 2019)

Debris Flow Model Sensitivity Analysis

The debris flow modelling was performed with the RAMMS::DEBRIS FLOW module version 1.7.20 (SLF/WSL, 2017a). This model was developed by Swiss research institutes and has since been used in science and practice (E3 and E6). Studies which implemented the debris flow module include Deubelbeiss and Graf (2013), Graf et al. (2019) and Schraml et al. (2015).

The overall approach to model debris flow uncertainty included repeated simulations with varying parameters. This type of method is commonly known as a sensitivity analysis. Its basic principle is to investigate the change in the model output when values of input parameters are systematically varied (Loucks et al., 2005; Uusitalo et al., 2015). Consequently, a variety of scenarios are constructed and separately tested through repeated model runs. The value range tested per input parameter must be representative of the distribution of the parameter in reality. In the most simple form of a sensitivity analysis, only one input parameter is changed at a time. However, it is also possible to combine multiple parameters to investigate their collective effect (Loucks et al., 2005). If a change in an input parameter results in a major difference in the model output, the model is sensitive to that parameter. A sensitivity analysis can be seen as a simplified version of the frequently used Monte Carlo simulations (Uusitalo et al., 2015). The study by Kunz et al. (2011a) presented in Chapter 2.9 served as a major source of inspiration for the overall approach to model the uncertainty data. They also conducted sensitivity analyses of numerical models for their uncertainty visualizations. Additionally, this type of analysis was also chosen by Schraml et al. (2015) to investigate the sensitivity of two debris flow models, one of which was RAMMS::DEBRIS FLOW, towards the erosion parameters and the expected event volume.

To implement the sensitivity analysis, the debris flow volume was changed between every simulation. The order of magnitude of the volumes were gathered from literature (see Volume Range [m^3] in Table 6). The indicated volumes were partitioned into ten volumes per location which were modelled separately. The volumes suggested by literature for the locations BON, LTB-T and LNK had to be extended in order to create meaningful spatial extents.

The following procedure was implemented to perform the debris flow modelling (see Figure 20 and the figures in Appendix G for accompanying screenshots). To start off the modelling workflow, a digital elevation model (DEM) was loaded into RAMMS in a .tif raster format. The DEM data used for this study was the swissALTI3D dataset by swisstopo (2019). This dataset is of high quality, which is important as the topography is a key parameter of the model and highly influences the spatial trajectory of a debris flow (Walser et al., 2014). A calculation domain (spatial extent of the simulation) as well as a hydrograph (release area of the debris flow) were drawn manually to initiate a simulation. The simulation resolution was set to 10m to keep computation times rather low. The friction parameters μ and ξ are important model inputs. ξ describes the viscous-turbulent friction, which is related the fluid parts of debris flows. It is recommended to use a value of 200m/s², which was kept constant throughout all debris flow simulations. μ defines the Dry-Coulomb type friction of the solid components of debris flows. Its values usually range between 0.05 and 0.4 (SLF/WSL, 2017b). To create simulations with a meaningful spatial extent, low friction values were necessary, so μ was set to 0.05 for all simulations. These erosion parameters also conform with the sensitivity analysis of RAMMS::DEBRIS FLOW performed by Schraml et al. (2015) during which past debris flow events were back-calculated. They found that in terms of the runout distance, μ has a larger effect than ξ, which was found to be responsible for the flow behaviour in the channel. Generally, they received best results when choosing values of μ between 0.07 and 0.11 and values of ξ ranging from 200 to 300 m/s². Lastly, the following hydrograph properties were used: The respective volume was entered, t1 (time of maximum discharge) was set to 10s and the flow velocity v was set to 5m/s. Based on these parameters the maximum discharge (Qmax [m³/s]) as well as the end time (t2 [s]) were calculated by the model.

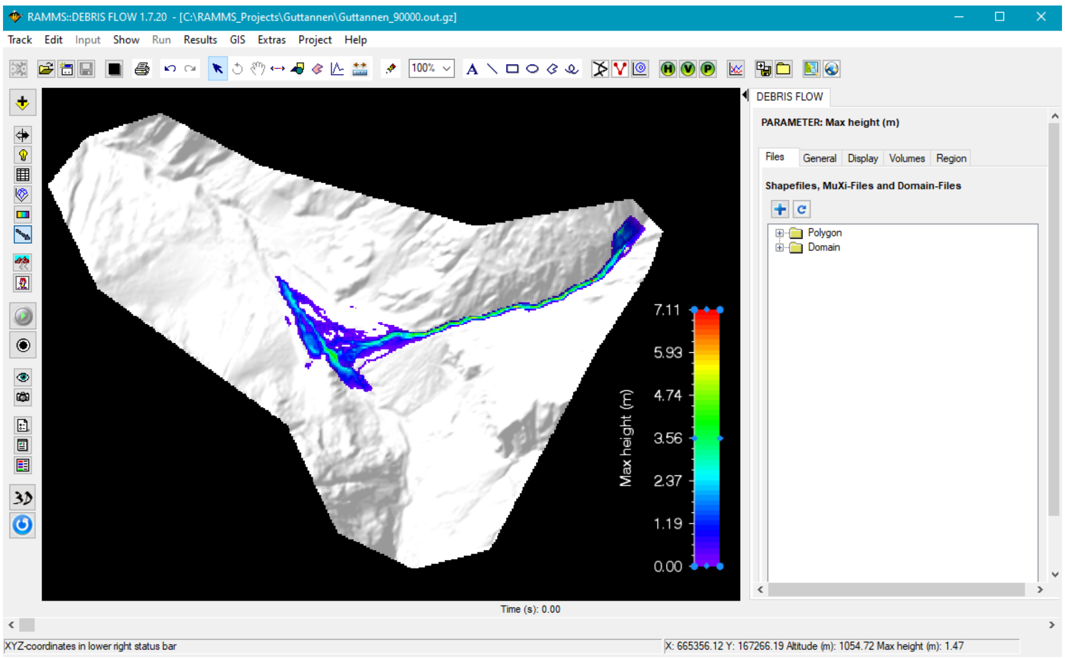


Figure 20: Interface of RAMMS::DEBRIS FLOW showing a simulation result for the location GUT with a volume of 90'000m³. The legend shows the maximum flow height [m].

Ten simulations with a varying volume parameter were prepared per location. These simulations were run with the RAMMS batch mode, which made the modelling procedure more efficient. The spatial extent of the maximum flow height of the modelled debris flow (.asc raster) was the data product of interest for the further processing.

Debris Flow Model Outcome Processing

The further processing of the debris flow simulations to arrive at the uncertainty information were performed in QGIS Desktop 3.16.4 with GRASS 7.8.5 (QGIS.org, 2020). A model description of the QGIS workflow is displayed in Figure 61 in Appendix H. The ten different maximum flow height raster images per location were loaded into QGIS and converted to binary masks. A value of 1 was assigned if the raster cell was affected by the debris flow in that specific simulation and a value of 0 was set if the raster cell remained unaffected. The masks were then added up. This resulted in a raster with values ranging from 0 (not affected in any simulation run) to 10 (affected in every simulation run). The resampling from a spatial resolution of 10m to 5m led to a smoothing of the uncertainty data with less abrupt categorical changes. The raster values were then inverted and rescaled to a value range of 0 (0% uncertainty) to 100 (100% uncertainty).

The uncertainty visualized in the map stimuli thus represents a likelihood that a certain location will be affected by a future debris flow event. This describes a measure of positional accuracy according to the typology of the SDTS (USGS, 1997). One could also interpret it as an attribute uncertainty regarding the debris flow volume. Furthermore, the uncertainty visualized in this study contains components of both aleatory (natural variability of future events) and epistemic uncertainty (caused by a lack of knowledge about the future) (Kunz et al., 2011b). Since the uncertainty was computed in the form of percentages, it represents a direct type of uncertainty based on the definition in Chapter 2.3.2 (Van Der Bles et al., 2019). As the data was created with a debris flow model, ontological uncertainty related to the model itself is likely present (Padilla et al., 2021a). The uncertainty is framed similarly as in Retchless and Brewer's (2016) study on displaying uncertainty in global temperature maps. A low uncertainty means that debris flow modelling was precise and that model results are confident that a house will be affected by a future event. A high uncertainty implies that confidence in the modelling results is lower meaning that it is difficult to say whether the house will actually be affected or not.

In a last step, the uncertainty layers were styled as described in Chapter 3.4.1.2. The map stimuli were created with map layouts in QGIS and exported as .jpeg images at a spatial resolution of 300dpi.

3.4.1.2. Uncertainty Visualization

The variety of uncertainty visualization methods, which have been developed during the past decades, is enormous. Consequently, MacEachren et al. (2005) criticized the discipline of uncertainty visualization for not providing enough empirical evidence on the effectiveness of the different methods. However, only a small selection of these methods can be tested at once in an empirical study. The methods chosen for this study represent intrinsic ones, which means that the uncertainty representation is shown on the data itself through the alteration of its appearance (Gershon, 1998). Intrinsic methods have been suggested to be more suitable for a non-expert audience as data and uncertainty information are not displayed separately (Deitrick and Edsall, 2006). A similar observation was made by Slocum et al. (2003) in their study investigating different visualization methods for displaying the uncertainty in water balance model outputs. They found that while experts preferred extrinsic visualization methods, decision-makers liked intrinsic ones more, as it gave them a better overview of the situation. Additionally, it is proposed to use intrinsic methods to display quantitative information (Kinkeldey et al., 2014). Since the study at hand is conducted with a non-expert audience and the goal is to use the map displays for decision-making, intrinsic methods are seen as a suitable choice.

Based on a variety of studies including the ones discussed in more detail in Chapter 2.4.5, colour value and colour hue – two methods making use of the visual variable of colour – were chosen for the study. They represent a type of pseudo-colouring which involves the matching of data values with a colour scheme that varies one or multiple colour properties (Borland and Taylor, 2007; Ware, 2013).

Colour can be decomposed into its three components hue, value and saturation (Chesneau et al., 2005). Hue represents the wavelength of a colour on the electromagnetic spectrum and is generally referred to with colour names such as *red* and *yellow* (Brewer, 1994; Roth, 2017). Value describes the brightness or luminance of a colour (Chesneau et al., 2005). And saturation gives insight into how pure a hue is (Ware, 2013). A highly saturated red seems strong while a low-saturation red is a pastel tone. Colour is extremely popular in the field of cartography as it can represent large amounts of data while preventing map readers from interpreting the information incorrectly (Chesneau et al., 2005; Sterba and Blaha, 2015). Furthermore, colour makes patterns otherwise hidden in data easy to grasp (Brewer, 1994) and enhances the visual attractiveness of map displays (Sterba and Blaha, 2015). The visual variable of colour was classified as an *undoubted (attention) guiding attribute* by Wolfe and Horowitz (2004). Additionally, colour is the most common method for visualization in risk maps (Chesneau et al., 2005).

The two methods to display uncertainty – colour value and colour hue – are now discussed in more detail before the development of the colour schemes for the map stimuli is explained.

Uncertainty Visualization Method 1: Colour Value

The visual variable of colour value is selective and non-associative and thus suitable for depicting ordinal data. However, colour value is frequently used to depict numerical data in choropleth maps (Roth, 2017). In MacEachren et al.'s (2012) study investigating whether visual variables are intuitive and uncertainty evoking, colour value reached the third place after fuzziness and location. In their study, colour value followed the *dark is more* principle. This means that certain data was displayed darker and the value became continuously brighter the more uncertain the data was.

Kubicek and Sasinka (2011) specifically investigated the intuitiveness of using colour value to depict uncertainty in soil depth data. A majority of their participants indicated that they interpreted lighter values to represent more uncertainty, which is in line with the *dark is more* principle. Additionally, participants with this opinion were also quicker to decide compared to ones who interpreted lighter values to show more certain data. Trau and Hurni (2007) conducted an analysis to investigate the suitability of various visual variables and methods to display uncertainty in hazard maps. It showed that colour value is suitable for univariate (adjacent) and bivariate (coincident) displays as well as specifically for the application in hazard maps. Leitner and Buttenfield (2000) conducted a study during which participants performed a siting task for a park and an airport in a region where the land cover type was associated with uncertainty. They chose colour value, colour saturation and texture to encode the uncertainty information. The most correct decision outcomes were found with colour value. However, unlike the *dark is more* principle mentioned above, these results were recorded when light values represented certain data. In a study by Bisantz et al. (2009), participants were asked to rank data uncertainty represented through the following methods: colour value, colour hue, colour saturation and transparency. They found that colour value enabled participants to correctly rank and rate uncertainty data with two types of map backgrounds (grid or map). It can, therefore, be assumed that colour value is a suitable and intuitive way to display uncertainty from a visual semiotics perspective.

Uncertainty Visualization Method 2: Colour Hue

Colour hue represents a selective and associative visual variable. Due to its unordered nature, colour hue is generally suitable for nominal data (Drecki, 2002; Roth, 2017). However, Chesneau et al. (2005) points out that colour hue is quite frequently used to display ordered data in map-making. One quality of colour hue is that it makes use of one out of seven basic colour contrasts, namely the contrast of hue (Sterba and Blaha, 2015). This contrast can make differences in the data as well as between the data and the background more easily distinguishable (Ware, 2013).

Regarding the implementation of colour hue for uncertainty visualization, Trau and Hurni (2007) stated that colour hue is suitable for univariate (adjacent) displays and *suitable but impractical* for bivariate (coincident) ones. They suggested that colour hue is suitable for depicting uncertainty in intensity maps but unsuitable for hazard maps. The map displays created for this study are seen as univariate as the uncertainty data is tied to the spatial extent of a potential debris flow event and no other debris flow parameter is displayed. From this standpoint, colour hue can be seen as a suitable method. Trau and Hurni (2007) saw colour hue as unsuitable for depicting uncertainty in hazard maps as defined by the Swiss framework. Nevertheless, the method was tested for the map stimuli of this study which represent a midway between hazard indication and hazard maps according to the definition by Zimmermann et al. (2005) in Chapter 2.8.

Joslyn et al. (2007) conducted an uncertainty visualization study during which participants were asked to predict the direction and speed of wind and decide whether or not to issue a wind warning. They found that, when uncertainty was displayed with colour hue, fewer warnings were issued if uncertainty was high and more if uncertainty was low compared to scenarios where no uncertainty was depicted. They concluded from their results that displaying uncertainty through colour hue helped participants to make more realistic forecasts. Cheong et al. (2016) found that colour hue, implemented through a spectral colour scheme, outperformed colour value and texture in terms of decision accuracy. As mentioned above, they suspected that this result was caused by the higher contrasts between hues compared to other methods such as colour value. Another potential reason for this result is that participants were familiar with the spectral colour scheme because of its frequent use (Ware, 2013).

The first hue-based colour scheme which comes to mind is most likely the spectral or rainbow colour scheme. The potential reason for this is its role as the default colour scheme in a variety of visualization and modelling tools (Borland and Taylor, 2007). Borland and Taylor (2007) have found that half of the papers published in the IEEE Visualization conference proceedings between 2001 and 2005 applying pseudo-colouring used the rainbow colour scheme.

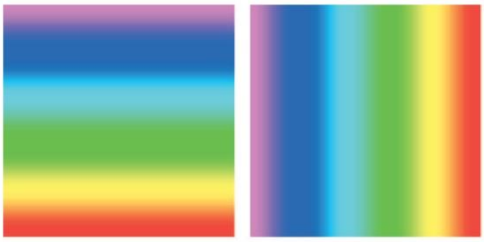


Figure 21: Examples of the rainbow colour scheme (Borland and Taylor, 2007).

The rainbow colour scheme contains an order in terms of the wavelength of the hues, however it lacks perceptual order (Borland and Taylor, 2007; Ware, 2013). This makes it challenging to interpret the order of the data represented by the colour scheme. Additionally, it contains abrupt boundaries between hues followed by sections during which hue appears constant (Borland and Taylor, 2007). These boundaries in the colour scheme hardly ever correspond with changes in the data (IPCC WGI Technical Support Unit, 2018). While inexistent transitions are implied at hue changes, the sections in between can obscure relevant patterns in the data (Rogowitz et al., 1996). For instance, the green part of the colour scheme in Figure 21 seems much larger compared to the bright blue section. Lastly, as the rainbow colour scheme simultaneously contains green and red, it cannot be correctly read by map readers with a protanomaly (IPCC WGI Technical Support Unit, 2018), which is the most common type of colour deficiency (Katsnelson, 2021). Nevertheless, colour hue is commonly implemented through some type of spectral colour scheme in uncertainty visualization methods (see Figure 22).

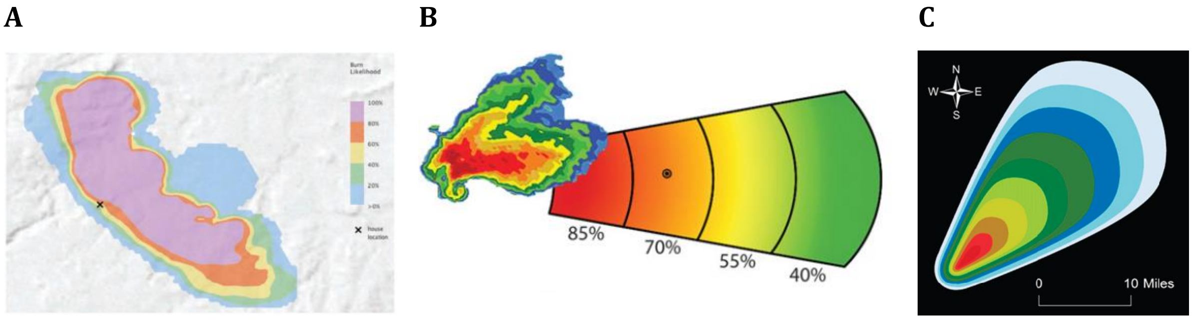


Figure 22: Three examples of spectral colour schemes implemented in uncertainty visualization maps: A – Cheong et al. (2016), B – Klockow-McClain et al. (2020), C – Ash et al. (2014).

Certainly the biggest challenge when implementing colour hue in a colour scheme is thus its lack of order. Bisantz et al. (2009) implemented hue through the selection of 12 colour hues and found that the results of the ranking task were very inconsistent between participants. Nevertheless, they see potential in hue-based colour schemes which do exhibit an order. During the literature review conducted for this thesis, no study making use of an ordered multi-hue colour scheme for the visualization of uncertainty was found. Therefore, this study provides first results on how colour hue can be used to depict uncertainty in a cartographically sound way.

Development of the Colour Schemes

The two visualization methods were termed after the manipulated visual variables during their introduction above. However, altering the colour hue often involves a change of colour value as well. Additionally, the term colour hue is frequently associated with the rainbow colour scheme. The two methods will be called single-hue for the colour value method and multi-hue for the colour hue method over the course of this thesis to create a clear picture and prevent biases.

The discussion on the rainbow colour scheme above has already shown that the choice of colours could lead to perceptual problems. It is, therefore, important that colour schemes are developed carefully in order for them to perform well. However, there is clearly no colour scheme which wins it all (Borland and Taylor, 2007). The choice of colour scheme strongly depends on the application context such as the task which will be solved with a specific map, the form in which the map is used or the cultural associations with different colours (Borland and Taylor, 2007; Sterba and Blaha, 2015). Therefore, different colour schemes were tested during the development of the map stimuli for this study. This chapter gives an overview of how the final colour schemes were prepared.

Brewer (1994) distinguished between four types of colour schemes: qualitative, binary, sequential and diverging. The colour schemes implemented for this study both represent sequential ones. This type of colour scheme is composed of classes which are ordered from low to high. Lower data values are displayed in darker colours and higher values in lighter colours. This again refers to the principle of *natural mapping* or *dark is more* which implies that dark colours represent low uncertainty, while light ones convey high uncertainty as implemented for instance by Bisantz et al. (2009) and Cheong et al. (2016). Importantly, Brewer (1994) suggested that multiple hues can be used in a sequential colour scheme if the brightness outweighs the difference in hue. The order of hues is an important aspect for the multi-hue colour scheme to enable the application of colour hue for visualizing quantitative data. Additionally, it is crucial that the step size in the colour scheme and the data is perceived as equal (Rogowitz et al., 1996).

The two colour schemes were strongly inspired by sequential colour schemes suggested by the ColorBewer 2.0 software (Brewer et al., 2013). These have been thoroughly tested and are widely used in the field of cartography. The usage of these colour schemes guaranteed that the results of this study were not negatively influenced by a lacking quality of the colour schemes. To keep results comparable, both schemes were based on the same colour hue overall. In terms of the choice of hue, different aspects were taken into consideration. As the map displays are used in natural hazard management, one would tend to use a red hue to imply the risk or danger posed by the debris flow, as suggested by the IPCC WGI Technical Support Unit (2018). Similarly, Zimmermann et al. (2005) proposed a red – orange – yellow colour scheme as an alternative to the present colours in Swiss hazard maps. However, it has also been found that different values of red are hard to distinguish, which could complicate the task (Katsnelson, 2021). Blue was excluded as Zimmermann et al. (2005) found that map readers misinterpreted it to signify a flooding event. In the end, green was chosen as the base hue for both colour schemes. This choice was inspired by one of the expert interviews, during which expert E1 mentioned that a green colour scheme is already in place for natural hazard intensity maps in Switzerland. Furthermore, olive colours were proposed for the mapping of debris flows by Kunz et al. (2011a).

A dark green colour was chosen as the starting point representing a high certainty for the single-hue colour scheme (see Figure 23A). This colour becomes continuously brighter, the higher the uncertainty is. Meanwhile, the hue is kept constant. In the multi-hue colour scheme (see Figure 23B), both colour value and hue are changed over the course of the colour scheme. Importantly, the starting point – in a dark green hue as well – and the end point – in a yellow hue – differ strongly regarding their brightness. This guarantees that the colour scheme is perceptually ordered despite the inclusion of multiple hues.

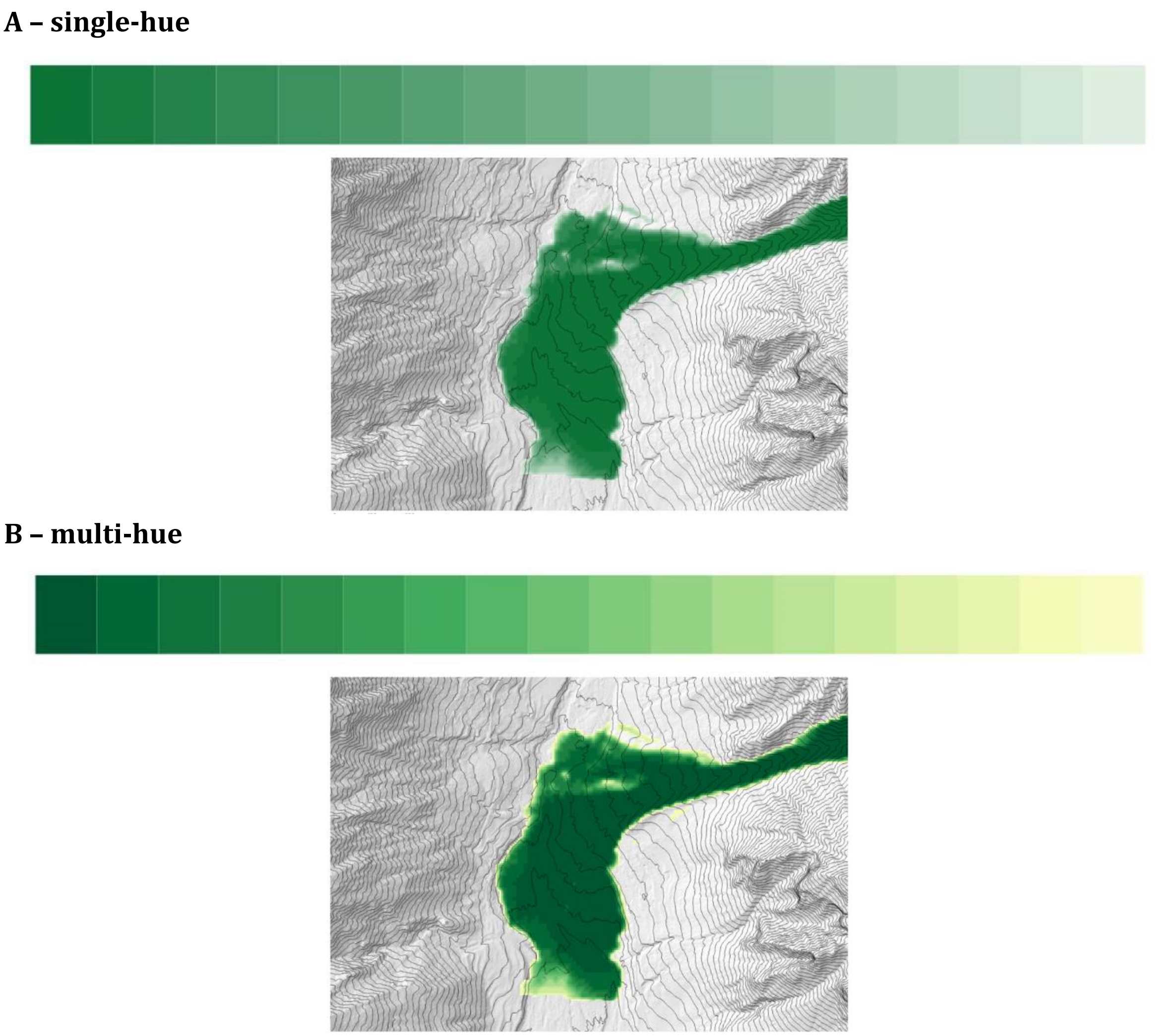


Figure 23: Colour schemes implemented in the map stimuli for the single-hue (A) and the multi-hue uncertainty visualization method (low uncertainties represented through colours at the left end of the colour schemes).

3.4.1.3. Uncertainty Communication

One major precondition for an uncertainty visualization to enable informed decision-making is the correct interpretation of the communicated uncertainty by its reader. As it has become clear in Chapter 2.6, most studies which directly compare numerical and verbal uncertainty communication are not map-based. For instance, Budescu et al. (2012) found that participants representing the general public systematically misinterpreted the text-based uncertainty regarding climate change. They concluded that a dual uncertainty scale combining numerical and verbal statements would be the best way to counteract this. The author of the study at hand, nevertheless, sees value in testing the two communication formats separately in a map-based setting. The map display contains additional information, for instance represented through the colour schemes, which was not present in text-based studies such as Doyle et al. (2014), Erev and Cohen (1990) and Olson and Budescu (1997). Previous map-based studies such as Ash et al. (2014), Klockow-McClain et al. (2020) and Miran et al. (2019) only compared how map displays including uncertainty communicated through numerical values (usually percentages) were interpreted compared to a deterministic scenario. Therefore, the study at hand stands out as it directly compares numerical and verbal uncertainty communication in a map-based context.

Uncertainty Communication Legends

As the results of studies discussed in Chapter 2.6 have shown, the interpretation of uncertainty in numerical and verbal form can exhibit large differences between people. The variety of expressions is highly variable, especially when it comes to verbal uncertainty (Budescu et al., 2012). Depending on their vocabulary, one individual might prefer a certain expression, while hardly using others (Wallsten and Budescu, 1995). During the development of the uncertainty legends it was of high importance to keep the individual influence of the author as small as possible. Therefore, the well established uncertainty communication framework implemented by the IPCC and presented by Mastrandrea et al. (2010) was used in this study. The framework helped to define the uncertainty labels as well as their placement along the value range. Wallsten and Budescu (1995) argued that the standardization of uncertainty terms and values has the potential to counteract the intra- and inter-individual differences in how uncertainty expressions are interpreted. However, Budescu et al. (2012 and 2009) have found that people struggle to correctly interpret the communicated uncertainties with this exact framework. Nevertheless, the IPCC likelihood scale is a valuable resource due to its frequent implementation. The usage of an established framework guaranteed that the uncertainty scale itself did not need to be tested in advance of this study. Additionally, this meant that findings of this study could be compared to previous results (e.g., Doyle et al. (2014)).

Table 7: Description of the official IPCC likelihood scale and its transfer to the uncertainty legend labels of the map stimuli.

IPCC Likelihood Scale		**Map Stimuli - Uncertainty Legend Label**	
Likelihood	**Term**	**Numerical**	**Verbal**
99-100%	Virtually certain	-	-
90-100%	Very likely	0%	Sehr wahrscheinlich
66-100%	Likely	33%	Wahrscheinlich
33-66%	About as likely as not	50%	Ebenso wahrscheinlich wie nicht
0-33%	Unlikely	66%	Unwahrscheinlich
0-10%	Very unlikely	90%	Sehr unwahrscheinlich
0-1%	Exceptionally unlikely	-	-

Table 7 gives an overview of the IPCC likelihood scale, the associated terms and how this framework was transferred to the uncertainty legend labels of the map stimuli. The percentage values were reversed for the numerical uncertainty legend labels as the stimuli described the uncertainty in an outcome instead of its likelihood as suggested by the IPCC scale. The official German translations of the IPCC likelihood terms were extracted from IPCC (2014). The only alteration made was changing the likelihood label between 33-66% – *about as likely as not* from *etwa ebenso wahrscheinlich wie nicht* as defined in IPCC (2014) to *ebenso wahrscheinlich wie nicht* (translating to *as likely as not*) in order to keep the label length shorter. Lastly, the most extreme IPCC likelihood levels were not implemented in the study. The 99-100% likelihood was already covered by the 0% uncertainty and the term *very likely*. The category 0-1% – *exceptionally unlikely* was completely excluded as the uncertainty data in the map stimuli was cut off at a level of 90% uncertainty to prevent confusion in the interpretation of higher values. The final map legends, which were implemented in the map stimuli, are shown in Figure 24.

3.4.1.4. Map Design

The overall goal of the remaining map design was to keep map complexity as low as possible and to focus on task relevant information to enable efficient map reading. An example of the result of the map design process is displayed in Figure 25. A further selection of map stimuli can be found in the Appendix P.

The base map was minimalistic to guarantee for a good contrast between the debris flow uncertainty data and the background. The hill shade, created from the swissALTI3D DEM, was the only layer in the base map. It was displayed with a transparency of 30% to keep the base map bright. This abstract base map was also chosen to prevent participants from recognizing the debris flow locations which could bias decision outcomes (Cheong et al., 2016). Therefore, no place names or other landscape features were displayed. Bisantz et al. (2009) showed that task performance did not significantly differ between two different base maps in their uncertainty visualization study. It was thus expected that this base map did not have an influence on the decision outcomes recorded in the study.

The contours were computed from the DEM as well. The distance between contours was set to 10m and they were displayed in black to ensure an appropriate contrast between the contours and both uncertainty colour schemes. Different symbols for the house locations were tested throughout the development of the map stimuli. The options differed in terms of their shape (cross vs. point), size and colour. In the end, a black point was used as it created a strong contrast from the map background and did not conflict with the colour schemes in terms of the suitability for map readers with colour blindness. A trade-off between its saliency and the amount of information covered by the symbol occurred to determine the size of the point.

The legend was parted into two sections. The symbols for the house location and the contours were shown at the top. The uncertainty legend explaining the colour scheme and the corresponding uncertainty values were placed below. The uncertainty labels differed between the two uncertainty communication groups.

The map scale was included as a help to judge the spatial extents of the map displays. Another reason why the map scale was included lies in the hypothesis that the distance between the riverbed and the house location could have an influence on decision-making.

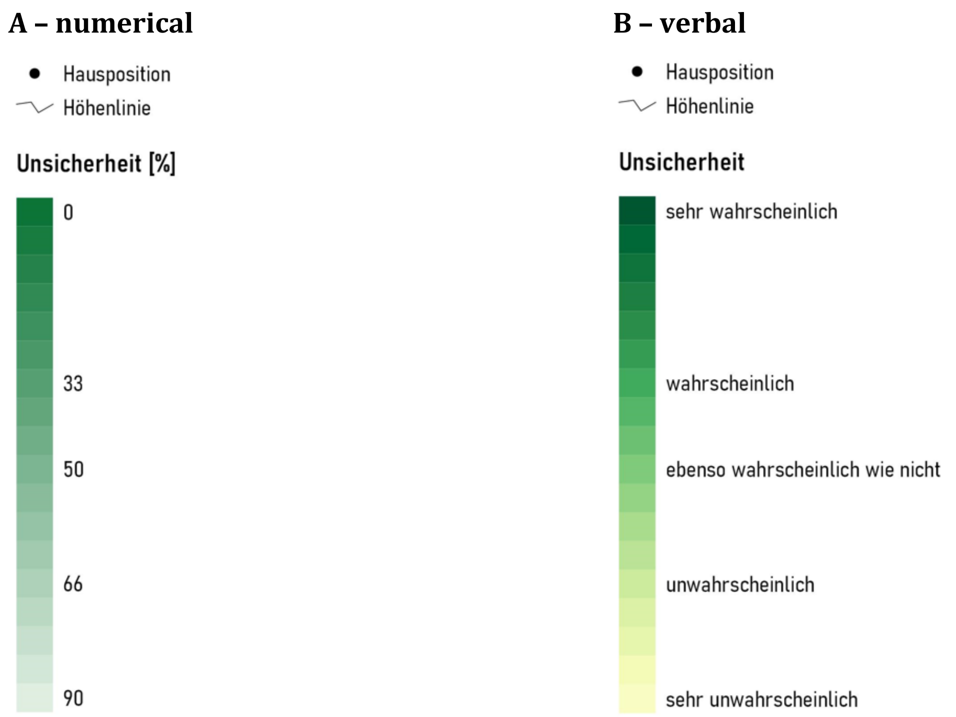


Figure 24: Map legends implemented in the stimuli showing an example of the single-hue (A – numerical) and the multi-hue (B – verbal) colour scheme for the two communication groups.

Linking back to the three types of visualizations by Buttenfield and Ganter (1990) presented in Chapter 2.1, the visualizations created for this study incorporate aspects of illustration and decision-making visualizations. On the one hand, the map displays can be seen as a rendered representation of a reality in the form of a GIS-based visualization for illustration. On the other hand, the purpose of the displays is clearly to enable informed decision-making with uncertainty.

Practitioner E5 emphasized in their interview that, although hazard maps are increasingly accessed digitally, the paper form of these maps is still important for field work. Cheong et al. (2016) support this standpoint under the consideration that static maps are often distributed among the general public especially in communicating spatial information for emergency planning. Due to these arguments, combined with the fact that Swiss hazard map are non-interactive at the moment, the map stimuli remained static.

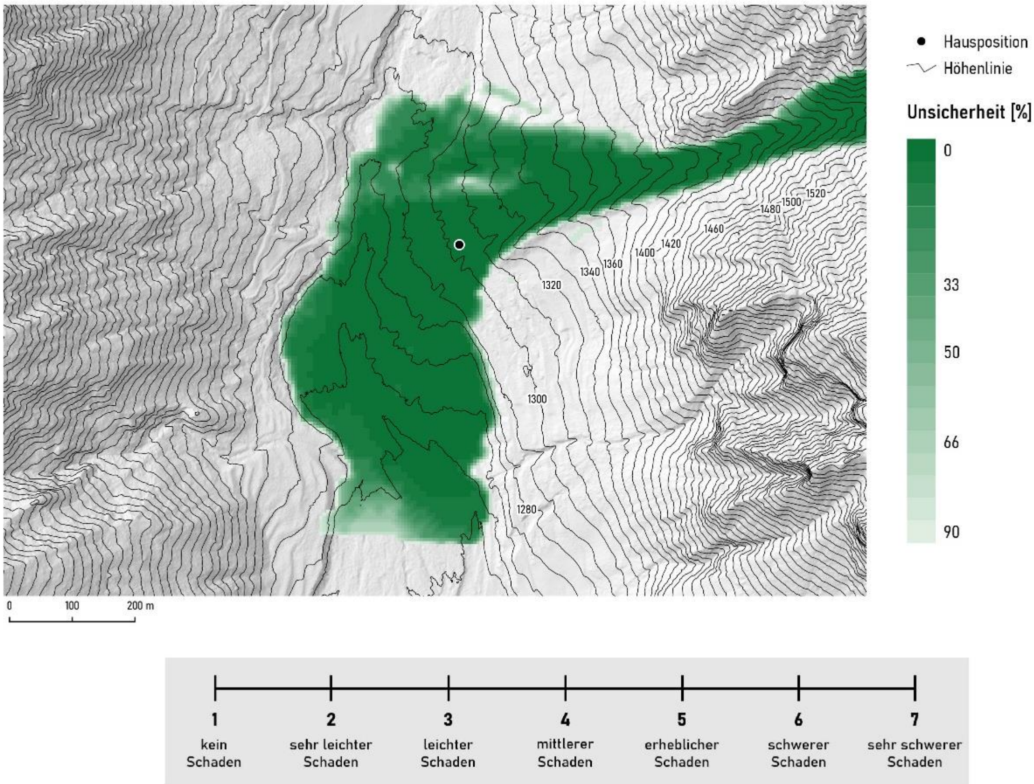


Figure 25: Example of a map stimuli at location Blenio (BLE).

In uncertainty depiction and general map making, it is important to be aware of the audience to which uncertainty or other spatial information is communicated (Fabrikant et al., 2010; Van Der Bles et al., 2019). Thus, it is important that the audience is able to read natural hazard visualizations and the accuracy of the map content properly (Zimmermann et al., 2005). The number of stakeholders in natural hazard management is large and includes politicians, scientists, insurances, spatial planners as well as the local population (Kunz et al., 2011a; Zimmermann et al., 2005). Inevitably, these stakeholders have varying experiences with natural hazards and map reading. Since the end product of natural hazard management such as a danger map is also directed at the general public, the map stimuli created for this study are targeted towards a non-expert audience with mixed map reading skills.

3.4.1.5. Saliency Analysis

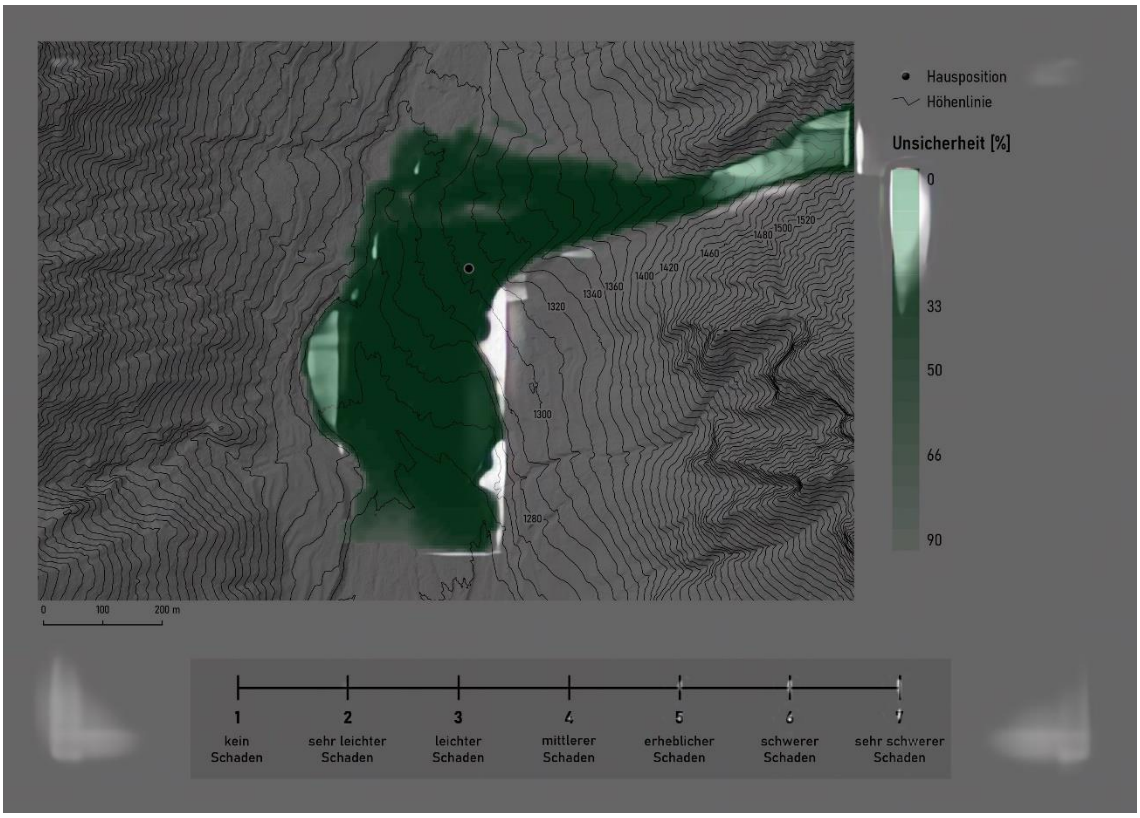
As elaborated in Chapter 2.2, attention strongly guides how humans perceive a visual scene. From a bottom-up perspective, the visual saliency characterized by properties such as colour or shape influences viewing patterns (Harold et al., 2016). Visual saliency is defined as the ability of an item to guide a viewer’s attention to itself. Thus, it is expected that humans first look at regions that are most different from their surroundings when there is no task associated with a display (Koehler et al., 2014). Saliency models have been developed to predict the way humans look at a visual scene. These models make use of the bottom-up processes involved in visual attention (Harold et al., 2016). They have been proven to successfully predict eye movements for the bottom-up perspective (Koehler et al., 2014). Yet, one aspect not taken into account in most saliency models are top-down cognitive processes. Previous knowledge or specific task instructions influence the way attention is guided, especially in real-world tasks (Fabrikant et al., 2010; Harold et al., 2016; Padilla et al., 2017). None the less, Fabrikant et al., (2010) see bottom-up saliency models as a useful tool to test map design choices in cartography.

From a map design perspective one can conclude that task-relevant information must exhibit a high saliency in order to support effective and efficient task solving. Specifically in the context of uncertainty visualization, McKenzie et al. (2016) state "that completeness of the visual display of uncertainty information (…) is less important than using visual cues that make the uncertainty information salient". Padilla et al. (2017) found that the visual saliency of the visualizations actually influenced decision outcomes in their study comparing different types of visualizations for the hurricane cone of uncertainty. However, they also call for caution as decision-makers could overestimate the importance of high-saliency features. Importantly though, one must always consider the combination of bottom-up and top-down processing when designing map displays (Harold et al., 2016).

The main purpose of the saliency analysis in this study was to guarantee that the results obtained in the experiment were stable. If saliency does not significantly differ between the two visualization methods and the two communication methods, it can be concluded that potential differences in damage estimates actually occurred due to the change in the visualization or communication method. Influences caused by a difference in saliency between map displays are then out of question. This approach was inspired by Fabrikant et al. (2010). The SaliencyToolbox – Version 2.3 (Walther and Koch, 2006) was used in MATLAB R2020a – Version 9.8 (MATLAB, 2020) to create saliency maps of the map stimuli. This tool is based on the saliency model developed by Itti et al. (1998). Walther and Koch (2006) state that "the model is meant to provide a first step to solving the chicken-and-egg problem of directing selective attention to object regions before objects are recognized". The three features which are taken into account to calculate the saliency map are colour, intensity and orientation (Walther and Koch, 2006). All three features were weighed equally to calculate the saliency maps.

Figure 26A & B show examples of the saliency maps for one single-hue and one multi-hue map stimulus. The highest values of saliency were recorded at the lower end of the colour schemes, at high elevation areas as well as the boundary of the debris flow extent. Additionally, the lower corners seemed to exhibit a higher saliency compared to their surrounding. However, this pattern disappeared when removing the orientation feature from the saliency map calculation. Most importantly, it is clear that the overall pattern of the saliency did not seem to differ strongly between the two visualization methods.

A – single-hue



B – multi-hue

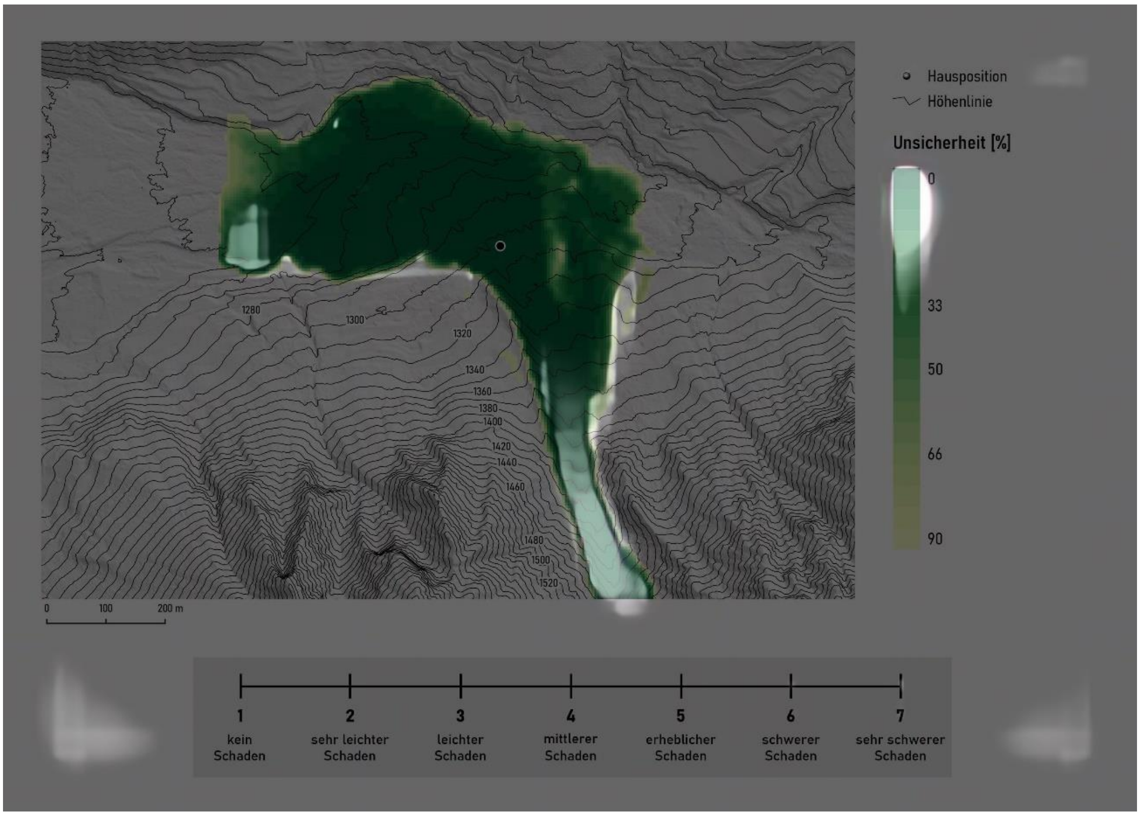


Figure 26: Saliency maps of tested map stimulus Nr. 1 (A - single-hue) and Nr. 2 (B - multi-hue). Bright areas in the saliency map represent areas, which are high in visual salience, while the dark areas are low in salience.

One concept closely related to saliency is visual clutter. Clutter describes a "state in which excess items [on a map display], or their representation or organization, lead to a degradation of performance at some task" (Rosenholtz et al., 2007). Clutter is influenced by the number of features on a map display. Yet, properties of these features such as their colour are also extremely relevant. Rosenholtz et al. (2007) developed two measures of visual clutter: the *Feature Congestion* and the *Subband Entropy*. The measure of Feature Congestion assumes that it is more challenging for an additional element to draw visual attention in a scene with higher visual clutter. Specifically, the measure has been found to effectively reflect a change in colour of a visual scene.

Table 8 shows a summary of the descriptive statistics for the Feature Congestion measures of the map stimuli. As the values were not normally distributed, Wilcoxon rank-sum tests were performed to investigate whether the Feature Congestion measures differed significantly between the visualization and the communication methods respectively. No significant difference in the means was found for the two visualization methods (p-value = 0.658). The values also did not significantly differ for the two communication methods (p-value = 0.334).

Table 8: Results of the Feature Congestion calculations for the different types of map stimuli.

	Normality	**Mean**	**Median**	**SD**	**Min**	**Max**
Overall	p-value = 0.002	4.68	4.60	0.82	3.27	6.14
Single-hue	p-value = 0.067	4.70	4.61	0.81	3.27	5.98
Multi-hue	p-value = 0.001	4.67	4.54	0.83	3.65	6.14
Numerical	p-value = 0.026	4.62	4.54	0.82	3.27	6.02
Verbal	p-value = 0.026	4.74	4.66	0.82	3.39	6.14

The Subband Entropy measures clutter through the degree of organization in a display by analysing its information content (Rosenholtz et al., 2007). A higher degree of organization, meaning that items can be grouped, leads to a lower value of visual clutter (Wilkening and Fabrikant, 2011).

The descriptive statistics of the calculated Subband Entropy measures are displayed in Table 9. The difference between the means of the two visualization methods was again found to be insignificant (p-value = 0.107). However, there seemed to be a significant difference in Subband Entropy for the two communication methods, as the p-value of 0.033 was lower than the significance level of 0.05. The effect of this significance was small to medium (r = -0.24) as the effect size r was below 0.3 (threshold for a medium effect) (Field et al., 2012). One likely reason for this difference could be the length of the uncertainty legend labels which represent high contrast areas due to the black font on a white background.

Table 9: Results of the Subband Entropy calculations for the different types of map stimuli.

	Normality	Mean	Median	SD	Min	Max
Overall	p-value = 0.017	3.60	3.61	0.19	3.27	3.94
Single-hue	p-value = 0.163	3.63	3.65	0.18	3.27	3.91
Multi-hue	p-value = 0.008	3.58	3.52	0.19	3.33	3.94
Numerical	p-value = 0.020	3.56	3.55	0.18	3.27	3.85
Verbal	p-value = 0.022	3.65	3.64	0.18	3.36	3.94

In conclusion, the map stimuli for the two uncertainty visualization and communication methods were found to be similar in terms of their saliency. Thus, the results of the study were not expected to be influenced by this characteristic of the stimuli.

3.4.2. Decision-Making Task

As described in Chapter 3.2 one of the measures recorded in this study is the decision-making outcome. Kubicek et al. (2012) distinguish between three types of information processing for tasks in uncertainty visualization studies. Firstly, search tasks are restricted to the localisation of a certain element on a map display, e.g., finding a point symbolization on a map. Recognition tasks include the matching of different elements in a map display, such as matching a certain hue on a colour scale to an uncertainty value in the legend. Lastly, in inference tasks, which are the most complicated type of task, participants need to combine their knowledge with the information presented on a map. Inference tasks are very common in uncertainty visualization studies. Even within the inference tasks a variety of different types can be found in literature. These range from simpler tasks such as ranking or classifying different areas on a map according to their associated uncertainty (Padilla et al., 2015; Retchless and Brewer, 2016) over deciding whether or not to take actions such as protection or evacuation (Cheong et al., 2016; Joslyn et al., 2007; Witt and Clegg, 2021) to self-reports on emotions such as fear (Miran et al., 2019). Damage estimations are another type of task popular in the context of uncertainty visualization in natural hazard predictions (Liu et al., 2017; Ruginski et al., 2016).

Likely the most important aspect of a decision-making task in an empirical study is its ecological validity. This property describes how closely an experimental task resembles tasks in real fields of application (Padilla et al., 2021b). Windschitl and Weber (1999) criticized the lack of ecological validity in the majority of decision-making studies, which use gambling tasks to infer on decision-making processes. It has also been argued that the use of artificial tasks could downplay how complex the context of a real-world decision is and lead to a low motivation for participants to make thorough decisions (Kinkeldey et al., 2017; Nadav-Greenberg and Joslyn, 2009). Despite the restricted possibilities in experimental settings, the goal of reaching a high ecological validity was kept in mind when designing the decision-making task. Therefore, common decisions made with the help of hazard maps were revisited. Hazard maps are a valuable tool in land use planning, planning of structural protection measures, construction of infrastructure and buildings and emergency planning (Chesneau et al., 2005; Raetzo et al., 2002; Zimmermann et al., 2005). Clearly, decisions in those fields are made by a variety of actors and are embedded in a complex context with consequences for stakeholders ranging from municipalities over insurance companies to the local population (Kunz et al., 2011a). Some common decision-making tasks mentioned by the experts during the interviews were decision on whether a parcel of land could be used for construction, if protection measures are necessary for a specific building or general spatial planning decisions (E1, E2, E5 and E6).

The implemented decision-making task was inspired and the expert interviews and studies such as Liu et al. (2017) and Ruginski et al. (2016). It was framed as follows: Participants were asked to estimate the damage caused by a debris flow at a specific house location indicated on the map considering the uncertainty of the location being affected by an event. They were presented with seven damage levels in the form of a Likert scale ranging from 1 – *no damage* to 7 – *very severe damage* (see Figure 25). They entered their decision outcome by clicking on the respective level on the Likert scale. More details on the Likert scale, its development and implementation are presented in the next subchapter.

The task is seen as complex enough to qualify as an inference task as proposed by Kubicek et al. (2012). Simultaneously, it does not exceed the limited expertise of the non-expert audience of the study. The decision task was purposely designed in a way that there was no correct answer as inspired by various studies (Deitrick, 2012; Deitrick and Edsall, 2006; Kübler et al., 2020). The decision accuracy was not considered in this study. The potential change in decision-making outcomes caused by the two uncertainty visualization or communication methods was much rather of interest.

3.4.2.1. Likert Scale

The concept of the Likert scale was first introduced by Likert (1932) in an aim to quantitatively measure social attitudes in the context of psychological studies. Joshi et al. (2015) defined an attitude as "preferential ways of behaving/reacting in a specific circumstance rooted in relatively enduring organization of belief and ideas (around an object, a subject or a concept) acquired through social interactions". Therefore, the scale enables the transfer of a subjective answer into objective results as well as a transition from a qualitative aspect to a quantitative measurement (Boone and Boone, 2012; Joshi et al., 2015). The Likert scale represents a so-called psychometric. The traditional scale included levels ranging from 1 – *strongly approve* – to 5 – *strongly disapprove* with the following levels in between: 2 – *approve*, 3 – *undecided* and 4 – *disapprove*. Thus, study participants were asked to indicate to what degree they concur with a certain statement, also called an item (Joshi et al., 2015). Later on, Likert scales were extended to measure other variables such as frequencies or similarities (Clark and Watson, 2019). Strictly speaking, a Likert scale is composed of multiple Likert-type items which investigate the same attitude and are combined to arrive at a single score (Boone and Boone, 2012). However, in research a single Likert-type item is frequently treated as a Likert scale (Albaum, 1997).

In the face of the strict definition, the Likert scale used in this thesis represents a Likert-type item. Nevertheless, it is referred to as a Likert scale in this work. The scale was designed as a unipolar type which does not have a *neutral* or *undecided* level (Jebb et al., 2021). As mentioned above, seven damage levels were included in the scale. Debates regarding the choice of the number of levels have been going on for decades. When designing a suitable scale, a trade-off must be made between offering enough levels so the attitude of a participant is represented, while preventing participants from being overwhelmed by the number of options. Clark and Watson (2019) found that the quality of a scale increased up until six levels were used. Joshi et al. (2015) argued that a scale with seven levels could outperform the original with five levels as it "provides more varieties of options which in turn increase the probability of meeting the objective reality of people". Convinced by their argument, a seven-level scale was implemented in this study.

A potential bias which could occur when implementing a Likert scale is that participants interpret the scale and its levels differently (Clark and Watson, 2019). To prevent this, the Likert scale was presented to participants during the introduction of the task in the pre-test questionnaire (see Appendix O). In addition to the scale itself, participants were provided with exemplary damages. This gave them a clear picture of the type of damage corresponding to the respective damage level. To develop the exemplary damages, the opinion of the interviewed experts were taken into account. The experts E1 and E2 emphasized that the most common damages caused by debris flows are damages to infrastructure such as streets and railways as well as property damage to buildings. Additionally, E3 mentioned that because debris flows contain solid materials, they can develop the destructive force to destroy a house. Another source of inspiration was a damage classification developed by Jakob et al. (2012) in their work on estimating building damage through an index of debris flow intensity. The developed Likert scale including the exemplary damages as displayed in Figure 27 was reviewed by the three experts from the scientific community. To implement the scale in the study, all labels were translated to German.

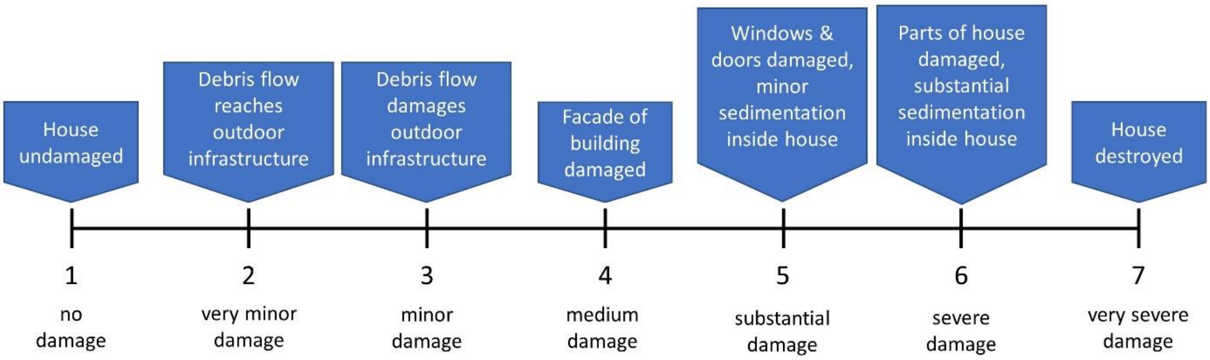


Figure 27: Likert scale used on the map stimuli with exemplary damages. The labels were translated to German for the experiment. Exemplary damages were removed on the actual map stimuli.

The Likert scale was also tested on two non-expert subjects to ensure that the order of the severity of damages as well as the exemplary damages are intuitive. The subjects were presented with the scale from 1 to 7 and were asked to assign the damage categories (*no damage* to *very severe damage*) as well as the exemplary damages (e.g., *Debris flow damages outdoor infrastructure*) to these levels. This task was performed with the German version of the scale. Both subjects were able to correctly assign the labels to the Likert scale.

3.4.3. Trial Planning

It has been shown that testing a large number of trials can be beneficial for the study outcome especially if the tested information is new to participants (Klockow-McClain et al., 2020). However, a high number of trials can lead to a long experiment duration and cause testing fatigue (Kübler et al., 2020). To prevent this, the trial planning of this study was prepared carefully.

The aim of the trial planning was to test house locations which fulfilled a variety of different conditions. The house locations were controlled in terms of the slope at the location as well as the distance between the house and the riverbed. Overall 40 trials were performed by each participant, which means that 20 locations were defined as each location was once tested in a single-hue stimuli and once in a multi-hue stimuli. Three slope categories (gentle, moderate, steep) were defined based on the slopes at the 10 debris flow locations. Similarly, the distance was split into three categories (short, medium, long) (for more information see Appendix I). The locations at the highest distance category were distributed around the boundary of the debris flow shape. Half of the locations were positioned just within the boundary and half of them were located outside the boundary. Importantly, no location was placed right on the boundary. This approach was inspired by studies investigating different visualization methods for the hurricane cone of uncertainty (Padilla et al., 2017; Ruginski et al., 2016) and served to capture a potential boundary effect in the results. The combinations of slope and distance categories were then systematically distributed across the 20 house locations. An overall number of 80 map displays were created as the map stimuli for the two communication groups differed in their legend.

Because each house location and debris flow site was used multiple times, the risk of a learning effect existed. This means that a participant could notice that they judged the same house location twice, once with the single-hue colour scheme and once with the multi-hue one. A common method to counteract this problem is to rotate the map displays as done for instance by Kübler et al. (2020) and McKenzie et al. (2016). Therefore, the map displays were systematically rotated by 90° during the map making process.

3.4.4. Eye Tracking & Experimental Design with Tobii

3.4.4.1. Basic Principles of Eye Tracking

The method of recording eye movements is often seen "as a window into internal cognitive processes" (Çöltekin et al., 2009). The main assumption in eye tracking studies is that humans move their gaze to those elements of a scene which they process at that moment (Just and Carpenter, 1976). The coordinates of the gaze point on which the eyes rest at a specific point in time are thus recorded (Ooms and Skarlatidou, 2018). Eye tracking can be seen as a way to quantitively measure the trajectory of visual attention (Montello et al., 2018). Because it is not based on self-report, eye tracking is considered to be an objective method (Brus et al., 2012; Li et al., 2010). Likely the only drawback of eye tracking is that the collected data can be extremely large in size and challenging to interpret (Ooms and Skarlatidou, 2018). Originally, the method was applied in reading research. However, thanks to the rising availability, the decreasing costs and the increasing computational power, eye tracking has spread across other fields of research (Irwin, 2004). In cartography, eye tracking has been used for data collection since the 1970s (Montello et al., 2018). It has since helped researchers to create efficient and understandable visualizations, also in the realm of uncertainty visualizations (Brus et al., 2012).

Eye movements can be classified into different types. The most important one are eye fixations which describe intervals of time during which the eyes rest on a specific element. The fixations are interrupted by saccades which are very rapid movements of the eye (Çöltekin et al., 2009; Galley et al., 2015). Based on the eye movement patterns, different metrics can be calculated. Some metrics which are of interest in eye tracking studies are the fixation duration, the number of fixations, the time to the first fixation and the number of saccades (Tobii Pro AB, 2014). Additionally, eye movements can be analysed spatially with the help of fixation coordinates and so-called areas of interest (AOI), which are discussed in more detail below.

3.4.4.2. Implementation in Tobii Pro Lab

The Tobii TX300 screen-based eye tracker was used to record the eye movements (more information in Chapter 3.6). Tobii Pro Lab is the software, which accompanies the eye tracker. It includes elaborate functionality to conduct psychological experiments applicable to this study. Therefore, the design of the main experiment procedure was implemented directly in Tobii Pro Lab – Version 1.171.34906 (x64) (Tobii AB, 2021). Two separate projects were created for the two communication groups and their respective map stimuli. Two timelines were built in each project, one for a trial task and one for the actual experiment (see Appendix J for accompanying figures). The stimuli were inserted into the timelines and separated by a break slide presented for 2 seconds. This slide showed a black cross with a white background and served as an anchor point between the stimuli (Tobii Technology, 2012). Instructions to start and end the trial were presented to the participants through text slides. To move from one map stimuli to the next, the participants were asked to click on one of the damage levels on the Likert scale. This click also recorded the decision outcome by registering the click at a specific damage level.

The order of the stimuli could induce a learning effect in repeated measure experiments (Wilkening and Fabrikant, 2011). Therefore, two types of randomization were implemented to change the order of the stimuli per participant. Firstly, the order of the two map stimuli sets – single-hue and multi-hue – was randomized. Some participants first solved tasks with the single-hue stimuli and some were first confronted with the multi-hue stimuli. Secondly, the order of the stimuli within each map stimuli set was randomized. This process is commonly referred to as counterbalancing (Martin, 2008).

Definition of Areas of Interest (AOIs)

To analyse the eye tracking data in a more systematic and statistical way, areas of interest (AOIs) were defined (Tobii Pro AB, 2014). AOIs describe parts of a stimulus which are of particular interest for a study (Brus et al., 2012; Korporaal et al., 2020). One eye tracking metric closely tied to the AOIs are visits. One visits starts with the first fixation recorded inside a specific AOI and ends with the last fixation registered in that area (Tobii AB, 2022a).

All relevant map elements such as the map display, the legend, the colour scheme and the Likert scale were covered by an AOI (see Figure 28). This resulted in a total of 9 map element AOIs and an additional 7 Likert scale AOIs (one per damage level). The purpose of the map element AOIs was to analyse the eye movement behaviour of participants and gain insight into the decision-making process. The Likert scale AOIs were created to extract the damage estimates during the processing of the data.

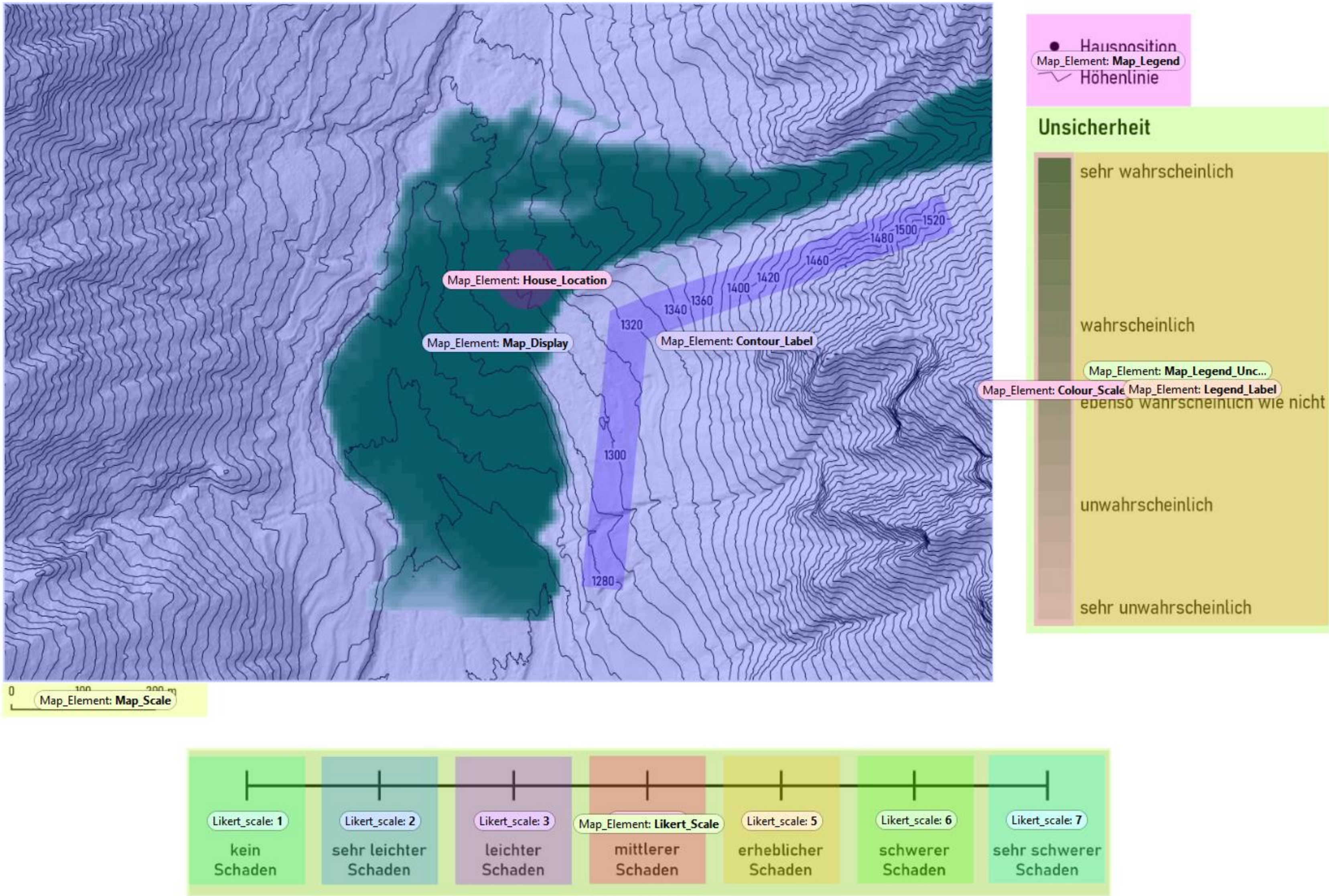


Figure 28: Division of map stimuli into map element AOIs and Likert scale AOIs based on a stimulus example of the verbal group.

AOI sizes were defined so the whole map element was contained within the shape. The only shapes that differed between the map stimuli were the ones of the uncertainty legend labels (numerical AOI was smaller compared to verbal AOI) and the contour labels which were adapted to the rotation of the map display. The width of the contour label AOI was informed by a preliminary analysis. This guaranteed that the AOI was wide enough to capture fixations but not too wide to negatively influence the eye tracking data. A similar approach was chosen to define the size of the house location AOI.

I-VT Filter

Extracting single fixations is an important step in eye movement processing with major effects on the analysis of the data. Fixations need to be separated and distinguished from saccades in a process called fixation identification (Salvucci and Goldberg, 2000). To do so, a variety of algorithms with varying underlying principles have been developed (for an overview see Salvucci and Goldberg (2000)). The algorithm offered in Tobii Pro Lab is a velocity-based I-VT filter (Olsen, 2012). Its main advantages are its simplicity and efficiency (Salvucci and Goldberg, 2000).

Tobii Pro Lab provides a default I-VT filter, which was developed based on the empirical study by Tobii Technology (2012). Since the software is applied in a variety of different studies, the default settings are suitable for a diverse range of applications (Tobii Technology, 2012). However, it is also possible to adapt these setting to customize the filter. One important parameter to define in the I-VT filter is the minimum fixation duration. The minimum fixation duration represents the lower limit of how long fixations are and serves as a threshold to discard fixations which are too short (Olsen, 2012). Defining the minimum fixation duration can be challenging and application dependent. While fixations are relatively short in reading, they last longer when processing an image or map (Tobii AB, 2022b). The default value in the I-VT filter is 60ms (Tobii Technology, 2012). Literature suggests other values such as 250ms (Galley et al., 2015), 200 to 400ms (Salvucci and Goldberg, 2000) or 50 to 500ms (Çöltekin et al., 2009). Based on these recommendations, as well as inspired by the map-based study by Fabrikant et al. (2010), the minimum fixation duration was set to 100ms.

Eye Tracking Data Export

One challenge tied to eye tracking is the vast amount of data which is produced (Li et al., 2010; Ooms and Skarlatidou, 2018). Instead of using the raw eye movement data, pre-processed metrics were exported to deal with the large datasets. Two types of data exports from Tobii Pro Lab were of interest in this study: the interval- and the AOI-based one. The software automatically creates intervals or automatic times of interest (TOIs) for screen-based experiments with stimuli. One TOI lasts as long as a specific stimuli is presented to the participant. The interval-based export contains one entry per TOI with the respective eye tracking metrics. The AOI-based export spatially aggregates the eye tracking data by AOI and TOI.

3.5. Post-Test

Debriefings after the experiment can be conducted in a variety of forms ranging from orally posed questions to questionnaires in written form (Martin, 2008). To perform both a qualitative and a quantitative analysis of the post-test answers, a structured questionnaire similar to the pre-test questionnaire was chosen in this study. Inspired by McKenzie et al. (2016) and Padilla et al. (2015) the main purpose of the post-test questionnaire was to gain insight into the decision-making strategies employed by the participants through self-report.

The post-test questionnaire was also implemented with LimeSurvey (LimeSurvey GmbH, 2021). It included some open-ended questions where participants described their decision-making strategies, the importance of different map elements and their preferences regarding the two colour schemes. They were also asked to rate colour schemes, task difficulty and the influence of the uncertainty information on their decision-making (see Appendix Q for the whole questionnaire). During the rating tasks, participants were encouraged to elaborate their answer through a comment. This helped to gain information on the reasoning behind their ratings.

3.6. Procedure & Experiment Environment

The study was conducted in the eye-movement lab at the University of Zurich, which is equipped with a Tobii TX300 screen-based eye tracker (data rate: 300Hz binocular, accuracy: 0.4°) (Tobii Technology, 2012). The eye tracker is connected to a 23'' monitor with a resolution of 1920 x 1080. The resolution was kept at this maximum value for the study to enable a highly resolved presentation of the stimuli. The eye-movement lab does not have any windows which ensured equal lighting conditions for all participants. As the study was conducted during the Sars Cov-2 pandemic, special measures had to be taken to ensure everyone's health. Participants and the experiment conductor were required to have a COVID-certificate (vaccinated or convalesced), to wear a mask and to practice social distancing. Additionally, surfaces at the workstation were disinfected after every participant and an air conditioning unit was running at all times.

After the participants were welcomed to the lab, they sat down at the work station (see Figure 29) and received overall information on the study. They started by reading and signing the consent form (see Appendix L), filled out the pre-test questionnaire and individually worked through the debris flow training session, the scenario introduction as well as the introduction to the task. The experiment conductor then introduced them to the eye tracking part of the experiment. Participants were moved into a comfortable position which enabled them to sit still.

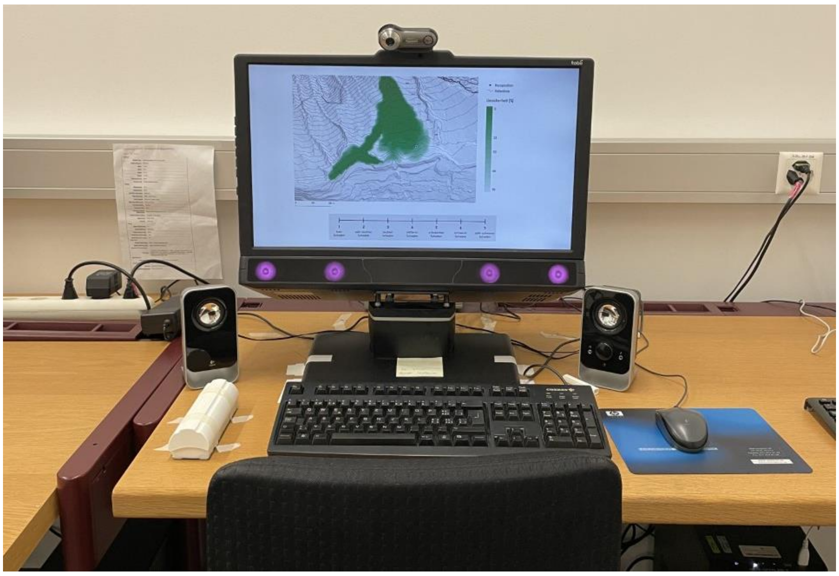


Figure 29: Setup of the workstation at the eye movement lab. The mouse was used to solve the task, a hand rest was offered for the left hand. The keyboard was used to fill out the questionnaires.

To guarantee for a high accuracy in the eye tracking data, the eye tracker was calibrated for each participant. This is necessary as every participant has a slightly different eye geometry (Tobii Pro AB, 2014). During the calibration a total of nine black target points were presented at different locations on a white background. The participant's task was to follow this point with their gaze as closely as possible. This procedure allowed the eye tracker to detect the participant's eye correctly to individually adapt the gaze estimation algorithm (Tobii Pro AB, 2014). After an initial calibration of the eye tracker, the participants were presented with two trial tasks to get familiar with the functionality of Tobii Pro Lab. The trial tasks were followed by another calibration which initiated the actual trials. After participants concluded the 40 tasks, they were asked to fill out the post-test questionnaire. All throughout the participants were offered short breaks and they had the possibility to ask questions if anything was unclear.

A draft version of the entire procedure was tested with two pilot participants. The aim of the pilot experiments was to test the procedure in a safe environment in advance of the actual study. Additionally, the pilot experiments were used to estimate the duration of one study trial. Both pilot participants took between 45 and 50 minutes to complete the study. Based on the feedback collected during the pilot experiments, certain questions in the pre- and post-test questionnaire were rearranged or rephrased. However, there was no need to change the overall setup of the study as well as the map stimuli. The preliminary data generated in the pilot experiments also ensured that the exported data formats were suitable for the planned data analysis.

3.7. Participants

The target audience of this study was a non-expert group of participants. This was motivated by the fact that most decision-makers using GIS or spatial information are actually not experts in that field (Hope and Hunter, 2007). Additionally, a goal of the recruiting was to find participants with a variety of different backgrounds. Uncertainty visualization studies so far were often conducted with audiences solely composed of students. The representativeness of these audiences for the general public was questioned by Kinkeldey et al. (2014) for instance. It was expected that participants had a basic knowledge of map reading but no in-depth experience with uncertainty visualization.

Participants were recruited via the author's personal and academic environment (for recruiting information see Appendix K). A total of 43 participant took part in the study: 25 were female and 18 were male. The participants were randomly assigned to either the numerical group (22 participants, gender ratio: 13 females – 9 males) and the verbal group (21 participants, gender ratio: 12 females – 9 males).

3.8. Data Analysis

The data analysis was split into a quantitative and a qualitative part. On the quantitative side, the decision outcome (damage estimate), the response time, the quantitative eye tracking data as well as some questions from the pre- and post-test questionnaire were analysed statistically. This analysis was conducted in RStudio – Version 1.3.1073 (RStudio Team, 2020). As mentioned above, the damage estimates were not directly contained in the Tobii Pro Lab data export but had to be derived. The damage estimates were extracted from the interval-based data by analysing the click metrics of the Likert scale AOI. On the qualitative side of the data analysis, the text answers from the questionnaires were downloaded from LimeSurvey and analysed manually by the author. Additionally, Tobii Pro Lab was used to qualitatively analyse heatmaps of the eye tracking data.

4. Results

4.1. Pre-Test

As mentioned above, 43 participants took part in the study. They were evenly distributed between the two communication groups (numerical group: 22 participants, verbal group: 21 participants). Overall the gender distribution in the two groups were almost equal with 13 female and 9 male participants in the numerical group, and 12 females and 9 males in the verbal group.

The overall mean age of participants was 30.05 years (SD: 9.56) with the majority of participants ranging from 20 to 30 years in age. The numerical group exhibited a slightly higher mean age of 31.23 years (SD: 10.61) compared to 28.81 years (SD: 8.4) in the verbal group. However, both groups shared the age range of 22 to 57 years.

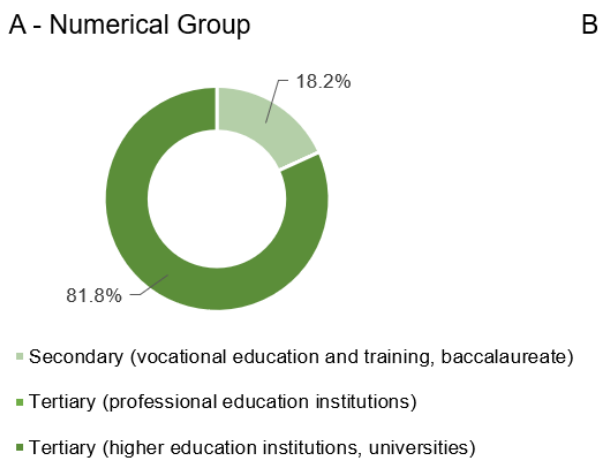


Figure 30: Relative distribution of maximum education for the numerical (A) and the verbal group (B).

Figure 30 shows the relative distribution of the maximum education for both groups of the study. Over 80% and 70% of participants had a degree from a university or a higher education institution in the numerical and the verbal group respectively, which indicates that the educational level is rather high. The pattern is quite similar between the two groups except for approximately 14% of people in the verbal group having a degree in professional education.

The results obtained from the self-assessment of the level of experience in a variety of topics are shown in Figure 31. For separate boxplots for the numerical and the verbal groups as well as the descriptive statistics see Figure 65 and Table 32 in the Appendix R, where additional results are provided. It is evident that participants rated their map reading skills rather high, while all other topics achieved medium to low ratings. Especially for the topic of debris flows, participants indicated low levels of experience. This shows that the subsequent thematic introduction in the pre-test was certainly valuable. Overall, the levels of experience deviated largely, which indicates that participants of various backgrounds took part in the study. Nevertheless, the low to medium experience results show that the participants represented a non-expert audience.

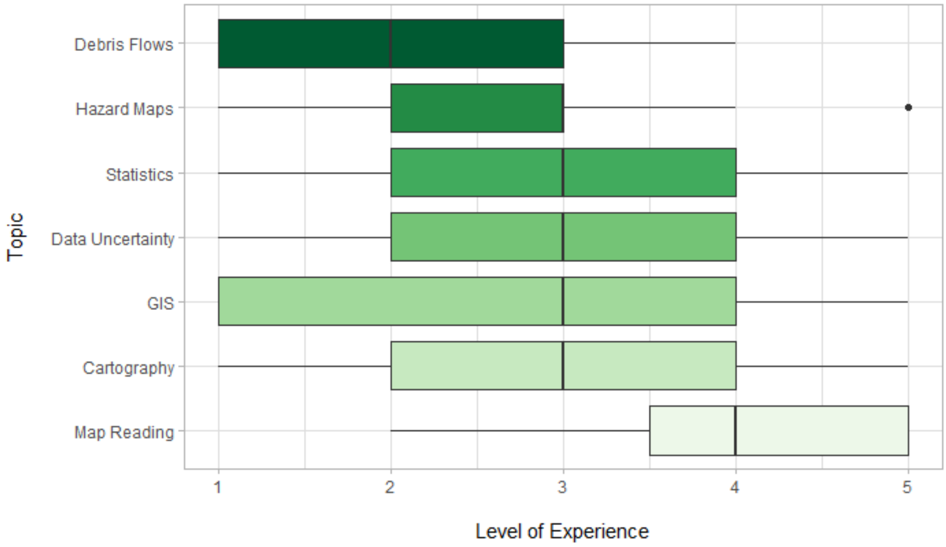


*Figure 31: Boxplot of self-assessments on the level of experience regarding different topics across all participants. The bold line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

In terms of how frequently participants use maps, the mean answer lay between a daily and a weekly usage. This answer is certainly related to the high level of experience in map reading. The relative distribution of answers to this question shows a similar pattern between the two communication groups with slightly more answers at lower map use frequencies in the verbal group (see Figure 32).

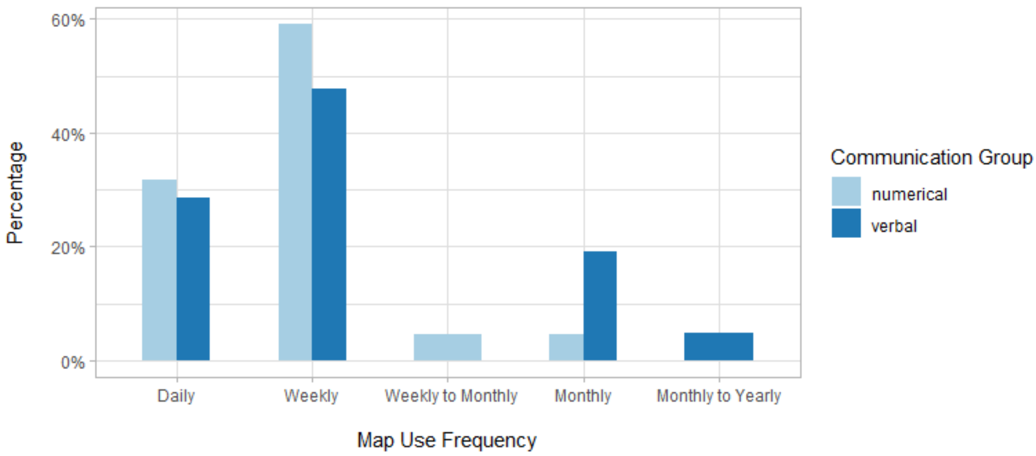


Figure 32: Relative distribution of map use frequency answers per communication group.

37% (n = 16) of participants indicated that they worked in one of the related fields. A total of eight out of these participants stated that they studied geography or GIS, six worked in the field of GIS and the remaining two worked in spatial planning and natural hazard related engineering. Yet, these participants were equally distributed across the two communication groups. In terms of experience in map creation 54.5% of participants in the numerical and 47.6% in the verbal group had already created a map themselves. This high percentage can certainly be traced back to the number of geography students and GIS trained participants taking part in the study. Only one participant in the verbal group had been personally affected by a debris flow in the form of a damage at their holiday house. Thus, a potential influence on damage estimates due to previous experience with the process, as suggested by Miran et al. (2019), can be neglected overall. Lastly, none of the participants indicated that they had previously been diagnosed with some type of colour vision deficiency.

4.2. RQ1: Uncertainty Visualization

The aim of the first research question is to investigate the influence of two uncertainty visualization methods on damage estimates, response time and decision-making strategies. To answer RQ1, the data was aggregated by the two visualization methods. This means that single-hue data contained results obtained from both communication groups, as the visualization method is a within-group variable of the study design. Information on potential interactions between the visualization method and the communication method can be found in Chapter 4.4. Because the visualization method describes a within-group variable with repeated measures of each participant, the single-hue and multi-hue results are dependent samples (Field et al., 2012). Please note that the results on decision-making strategies are presented in Chapter 4.6.1 for all the research questions.

4.2.1. Damage Estimate

Damage was estimated on a scale ranging from 1 – *no damage* to 7 – *very severe damage*. The overall mean damage estimate was 4.18, which represents a medium damage. The mean value did not vary strongly between the two visualization methods with an estimate of 4.10 for the single-hue stimuli and 4.26 for the multi-hue stimuli (see Table 10).

Table 10: Results of the damage estimate overall and per visualization method.

	Normality	**Mean**	**Median**	**SD**
Overall	p-value < 0.001	4.18	4	2.09
Single-hue	p-value < 0.001	4.10	4	2.10
Multi-hue	p-value < 0.001	4.26	4	2.09

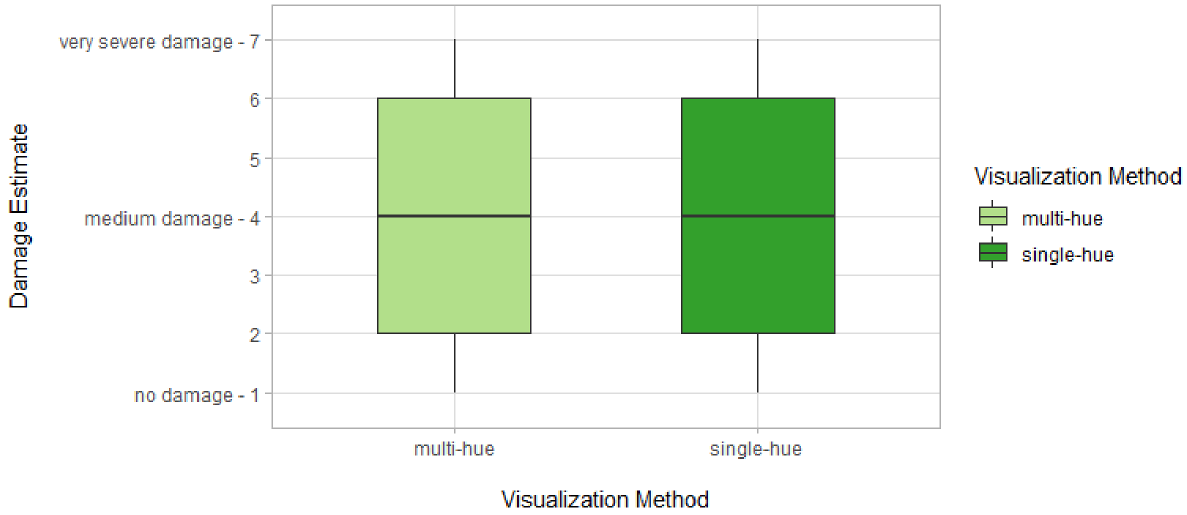


*Figure 33: Boxplot of the damage estimate per visualization method. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

Figure 33 shows that damage estimates likely did not vary between the visualization methods. For both visualization methods 50% of the estimates were registered between the levels 2 – *very minor damage* and 6 – *severe damage*. However, when looking at the relative distribution of the damage estimates in Figure 34 one can see that participants chose low damage estimates more frequently for single-hue stimuli and high damage estimates more often for multi-hue stimuli. The largest difference of 2.6% was found at the highest damage category (7 – *very severe damage*).

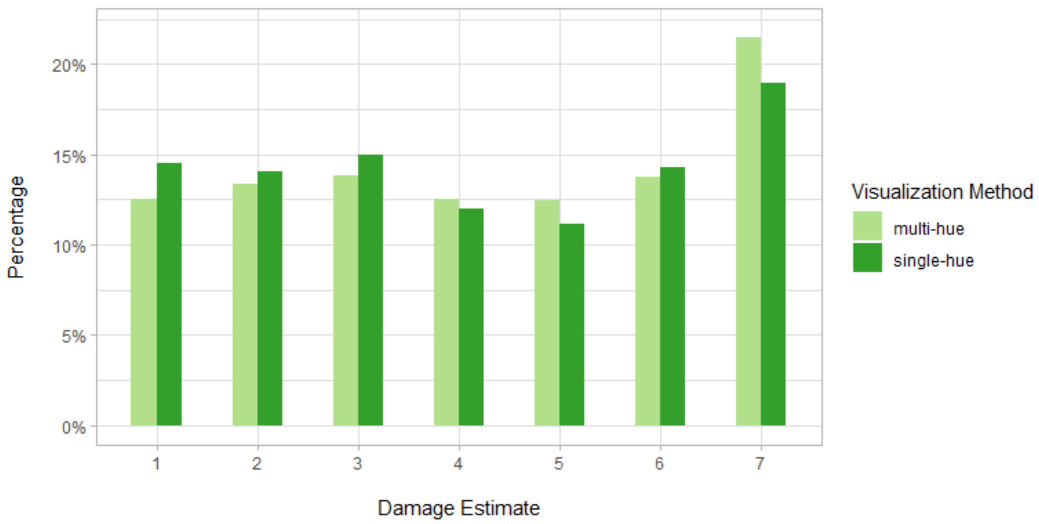


Figure 34: Relative distribution of the damage estimate per visualization method.

As indicated by the small p-values (< 0.05) in Table 10, the damage estimates for both visualization methods were not normally distributed. Therefore, a non-parametric test was chosen to compare the damage estimates per visualization method statistically. Because the visualization method is a within-group variable and the two samples are thus dependent, a Wilcoxon signed-rank test was performed (Field et al., 2012). The null hypothesis (H_0) of this test was that the means of the two samples were equal. The alternative hypothesis (H_A) stated that the means of the two samples differed. The p-value resulting from this test indicated that there was a difference in means (p-value < 0.001). This seems rather surprising after previously interpreting Figure 33.

It is important to know that even though the outcome of a statistical test can be significant, the measured effect of varying the variable may not be meaningful. This is why the effect size must also be investigated (Field et al., 2012). An effect size r of 0.1 is interpreted as a small effect, medium effects are observed at 0.3 and the effect is large if r is larger than 0.5 (Field et al., 2012). Effect sizes are generally reported as absolute values. For the Wilcoxon signed-rank test at hand an effect size of $r = 0.04$ was calculated. This means that the effect of the visualization method on the damage estimate was very small despite the significant result when investigating the statistical difference of means. As the significant result above came as a surprise, a dependent t-test was also performed even though the precondition of normally distributed samples was violated. Again, a significant difference was found ($t(859) = -5.01$, $p < 0.001$). The resulting effect size was $r = 0.17$, which again represents a rather small effect. It can thus be concluded that the visualization method led to a significant difference in damage estimates but this difference must be put into perspective with the small effect size.

4.2.2. Response Time

Overall, participants took 13.14 seconds (SD: 7.69s) to estimate the damage with the help of the tested map stimuli. They were slightly quicker when judging single-hue stimuli (mean: 13.02s, SD: 7.60s) compared to multi-hue ones (mean: 13.26s, SD: 7.77s).

Table 11: Results of the response time [s] overall and per visualization method.

	Normality	Mean	Median	SD	Min	Max
Overall	p-value < 0.001	13.14	11.43	7.69	1.90	67.13
Single-hue	p-value < 0.001	13.02	11.35	7.60	1.90	67.13
Multi-hue	p-value < 0.001	13.26	11.63	7.77	2.02	56.17

What stands out in Table 11 are the broad value ranges from approximately 2 seconds to over 1 minute. The upper range outliers can also be clearly identified in the boxplots in Figure 35. Yet, since those high response times were recorded for both visualization methods, there was no need to filter them out.

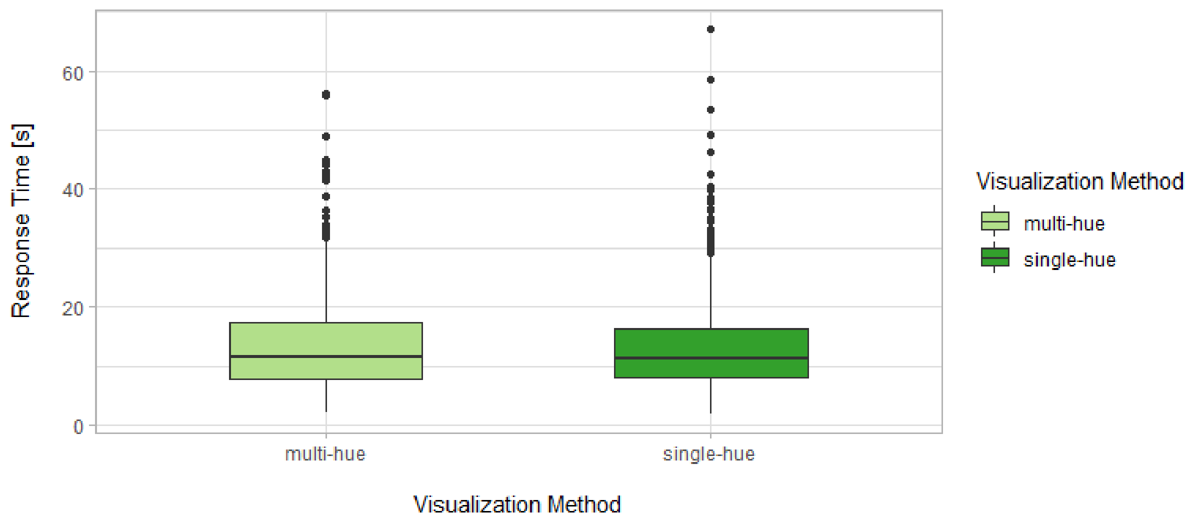


*Figure 35: Boxplot of the response time [s] per visualization method. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

Again the potential difference in the response time for the two visualization methods was analysed statistically. A Wilcoxon signed-rank test was performed as response time did not follow a normal distribution (see Normality column in Table 11). The test resulted in a p-value of 0.387, which is larger than the confidence level α of 0.05. Therefore, the null hypothesis was failed to be rejected and the response time did not significantly differ between to the visualization methods.

4.2.3. Eye Tracking

Two different eye tracking metrics were investigated in detail. The first of them is the mean total duration of fixations per AOI. This metrics describes the amount of time [s] participants spent looking at a specific AOI per trial. The second metric discussed is the mean number of visits detected per AOI. The number of visits reflects how often the participants' gaze entered a specific AOI, hence the number of times the AOI was visited (Tobii Pro AB, 2014). This mirrors the overall viewing pattern independently of the time spent in one AOI. A high number of visits overall means that participants moved their gaze more frequently between the different AOIs. This is a pattern which cannot be discovered solely by analysing at the duration of fixations.

Figure 36 shows that participants spent most of the time looking at the map display, followed by the house location and the Likert scale. This gives a first impression of which map elements seemed to be relevant to solve the task. Interestingly, the differences in total duration of fixations between the two visualizations methods seemed to be minor. Participants spent slightly more time on the map display for single-hue stimuli. However, the amount of time spent looking at the house location and the contour labels which were both contained by the map display were higher for the multi-hue stimuli. Potential differences between total duration of fixations per AOI for the two types of visualizations methods were investigated statistically through dependent t-tests or Wilcoxon signed-rank tests depending on the distribution of the data. The results in Table 12 show that significant differences were found for the uncertainty label and consequently for the uncertainty legend as a whole. The time spent in those AOIs was generally longer for the multi-hue stimuli. Yet, the low effect sizes need to be taken into account when interpreting these results.

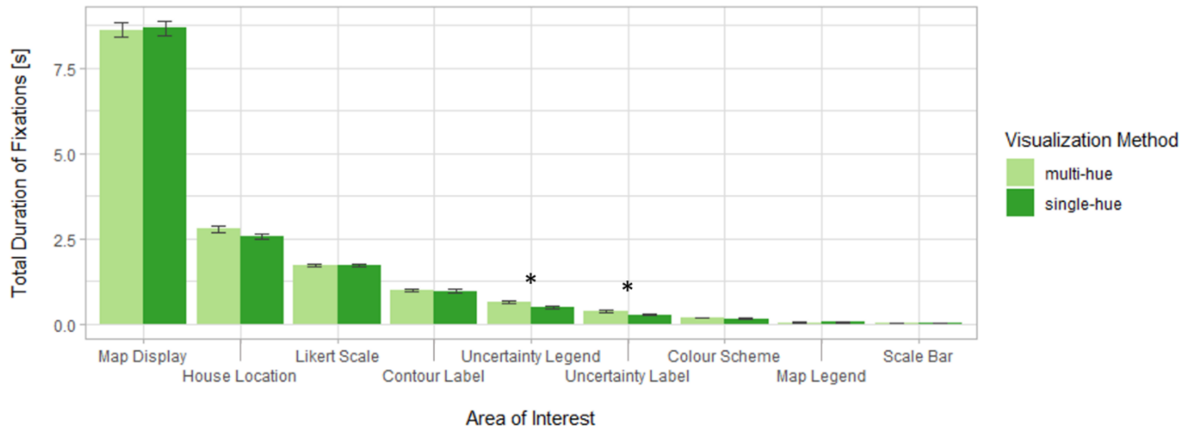


*Figure 36: Mean total duration of fixations [s] per AOI and visualization method. Error bars indicate +/- 1 standard error. The * symbol indicates statistically significant differences between the visualization methods.*

The strong effect of the map display could of course be caused solely by its large size. For further investigation, the mean total duration of fixations per AOI were normalized by their respective area. These results drew a rather different picture: The house location (one of the smallest AOIs) registered most of the fixation duration per pixel, while all other AOIs seemed to have been barely looked at. Additionally, participants spent more time on the house location for the multi-hue stimuli. The results gained from Figure 36 and Figure 37 must certainly be combined to create a complete picture of how participants spent their time looking at the maps.

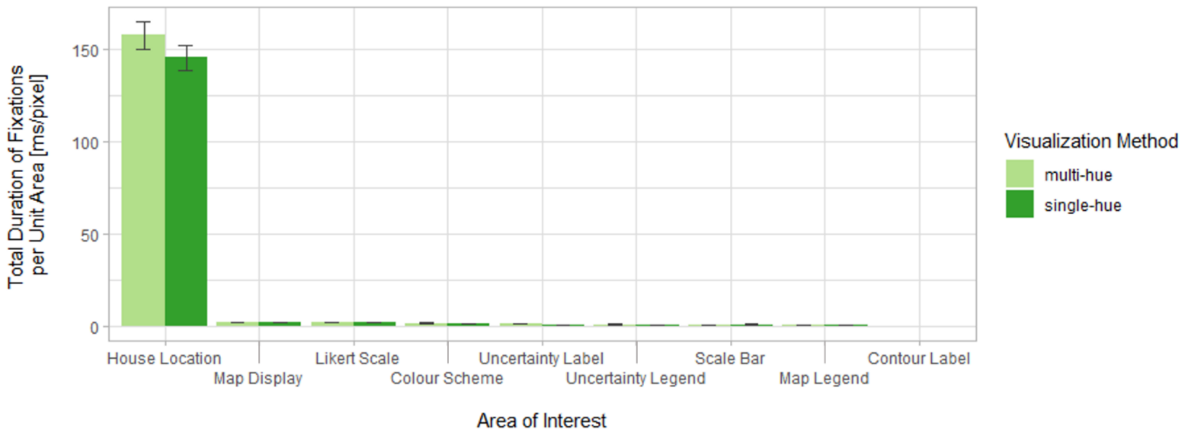


Figure 37: Mean total duration of fixations [ms/pixel] per AOI and visualization method, normalized by the AOI area. Error bars indicate +/- 1 standard error.

Table 12: Statistical results of the comparison of means for two eye tracking metrics between the visualization methods for each AOI. Italic entries symbolize that the difference was significant and, therefore, the effect size r is included.

AOI	Total Duration of Fixations	Number of Visits
Colour Scheme	p-value = 0.213	p-value = 0.073
Contour Label	p-value = 0.744	p-value = 0.264
House Location	p-value = 0.079	p-value = 0.624
Likert Scale	p-value = 0.342	p-value = 0.115
Map Display	p-value = 0.641	p-value = 0.236
Map Legend	p-value = 0.488	p-value = 0.820
Uncertainty Label	*p-value < 0.001 (r = 0.095)*	*p-value < 0.001 (r = 0.104)*
Uncertainty Legend	*p-value = 0.001 (r = 0.078)*	*p-value < 0.001 (r = 0.084)*
Scale Bar	p-value = 0.768	p-value = 0.295

On average a total of 14.25 visits (SD: 7.65) were recorded within the AOIs with single-hue stimuli. Comparatively, the mean number of visits on multi-hue rounded up to 14.95 visits (SD: 8.53). The rather high standard deviations could either be caused by interindividual differences in the employed strategies or because the number of visits necessary to arrive at a decision continuously decreased the more trials participants had performed.

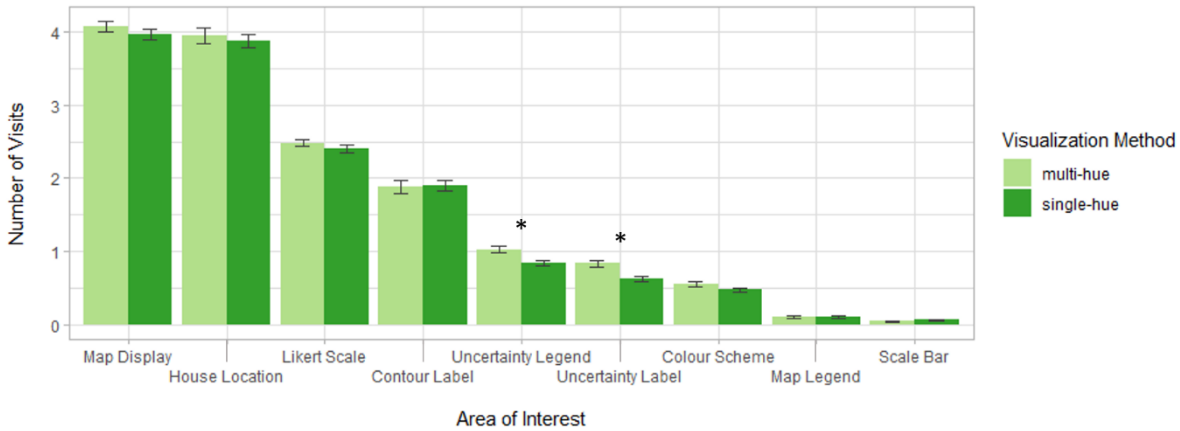


Figure 38: Mean number of visits per AOI and visualization method. Error bars indicate +/- 1 standard error.

The overall ranking of AOIs in terms of the number of visits was the same as observed above. However, the pattern drew a different picture as the map display did not dominate the ranking (see Figure 38). The mean number of visits shows that participants' gaze entered the map display, the house location and the Likert scale most often. In terms of the decision-making process, this pattern indicates that participants moved their gaze three to four times between the house location and the Likert scale in order to reach their decision, while hardly visiting the uncertainty legend and the colour scheme. The patterns of the number of visits were very similar for the two visualization methods. Significant differences were again found for the uncertainty label and the uncertainty legend as a whole. More visits were recorded for multi-hue stimuli compared to single-hue ones in these AOIs. Yet, the effect sizes of these significant differences were rather low (see Table 12).

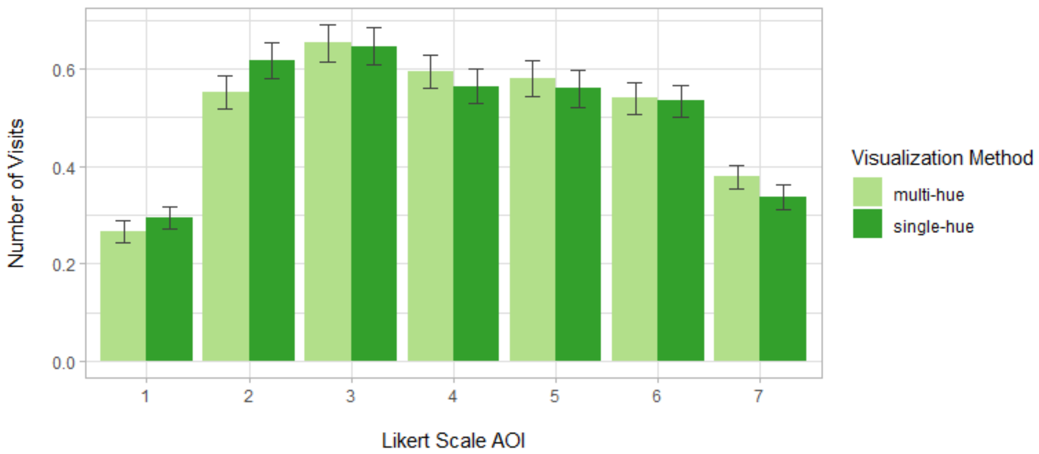


Figure 39: Mean number of visits per Likert scale AOI and visualization method. Error bars indicate +/- 1 standard error.

Figure 39 shows the number of visits detected at the different damage levels of the Likert scale. It clearly exhibits a similar pattern for multi-hue and single-hue which reflects the results of the damage estimates discussed above. However, it is interesting to compare this figure to the relative distribution of damage estimates in Figure 34. Evidently, participants made surprisingly few visits in the Likert scale level 1 – *no damage* and 7 – *very severe damage* compared to how often these levels were chosen in the decisions. The slight differences in the number of visits per Likert scale AOI between the visualization methods accurately reflect the differences in the relative distribution of damage estimates with the exception of level 3 – *minor damage* and level 6 – *severe damage*.

All the observations made from these plots are also reflected in the heatmaps (Figure 66) in the Appendix R. These maps show the relative time spent in a specific area of the stimuli (Tobii Pro AB, 2014). These maps qualitatively show that fixations were mainly detected in the map display. Additionally, the Likert scale pattern described above is clearly reflected. The heatmaps also qualitatively indicate that more time was spent looking at the uncertainty label for the multi-hue stimuli overall and especially for the verbal group. One might also notice that the viewing patterns on the map displays were slightly different for the two visualization methods, which was caused by the different orientation of the maps. However, it is clear that for both types of stimuli, the house location attracted a considerable amount of attention.

4.2.4. Post-Test

Two preference task in the post-test questionnaire specifically related to the visualization methods. Participants were asked to rate the suitability of the two colour schemes for their decision-making from a scale of 1 – *not helpful* to 5 – *very helpful*. Additionally, Sterba and Blaha (2015) argued that colour can contribute significantly to the visual attractiveness of a map. Therefore, participants also judged how much they liked the two visualization methods in terms of their aesthetics from a scale of 1 – *do not like at all* to 5 – *like very much*.

37 out of the 43 participants preferred the multi-hue method in terms of the suitability for decision-making (see Figure 40A). They assigned a mean rating of 4.58 (SD: 0.70) to the multi-hue stimuli. The single-hue method only achieved a mean score of 3.44 (SD: 0.91). The six participants who preferred the single-hue method were evenly distributed between the two communication groups. To study the difference in preference scores statistically, a Wilcoxon signed-rank test was performed as the scores for both visualization methods were not normally distributed (single-hue p-value < 0.001, multi-hue p-value < 0.001). The test showed that the difference was indeed significant (p-value < 0.001) and the effect was medium to large (r = 0.46).

Participants were also asked to indicate what aspects of the two colour schemes they liked regarding their suitability for decision-making. Three participants appreciated that the single-hue method was uncertainty evoking, an argument also made by MacEachren et al. (2012), and two participants thought the colour scheme was intuitive. However, almost half of the participants mentioned that the different brightness levels, especially the very bright values, were hard to distinguish. They also criticized that the low contrast between the debris flow extent and the base map made the decision-making challenging. For the multi-hue colour scheme 25 participants liked that the different colour levels were more clear. Additionally, 10 participants mentioned that the yellow hues used for high uncertainty values supported their decision-making for locations in high uncertainty regions. On the contrary, four participants stated that they perceived the high uncertainty values as too salient. A complete overview of all the positive and negative aspects can be seen in Table 33 in the Appendix.

A very similar picture was drawn in terms of the preference regarding the aesthetics. Again only six participants (three per communication group) preferred the single-hue method (see Figure 40B). This method reached a mean score of 3.05 (SD: 1.00). 36 participants preferred the multi-hue method which received a mean score of 4.19 (SD: 1.01). As one can see in Figure 40B, there was one participant, which rated both methods equally. As the samples were not normally distributed (single-hue p-value < 0.001, multi-hue p-value < 0.001), a Wilcoxon signed-rank test was conducted to investigate the difference in the preference scores. The resulting p-value was below 0.001, which points towards a statistically significant difference. The effect of the visualization method on the preference in terms of the aesthetics was medium to large (r = 0.43).

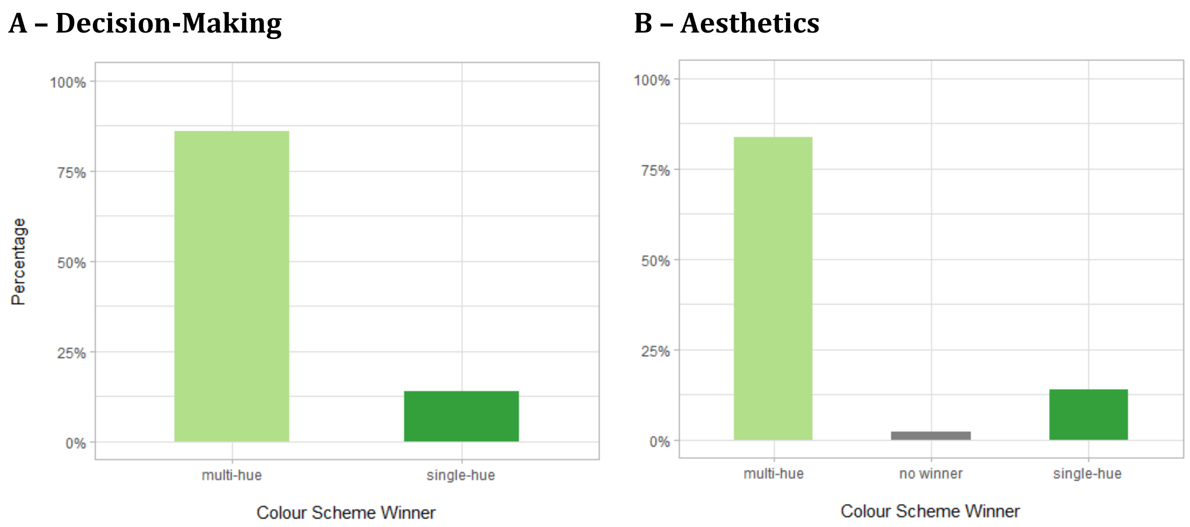


Figure 40: Results of preference tasks on the visualization methods regarding the suitability for decision-making (A) and the aesthetics (B). The winner represents the visualization method with the higher score per participant.

When asked for reasons why participants liked or disliked the aesthetics of the two colour schemes, the following reasons were mentioned: Participants liked that the single-hue colour scheme seemed calm and intuitive to read. However, others thought the colour scheme was bland. They mentioned that the multi-hue colour scheme was aesthetic and easy to interpret. They also appreciated the high saliency and the contrast between the hues. One negative aspect mentioned by five participants for both colour schemes was that they did not interpret green as danger or risk evoking. These participants would have preferred another colour hue as the basis of the colour schemes. A comprehensive overview of all the positive and negative aspects mentioned can again be found in the Appendix (see Table 34).

Overall, it seemed that the multi-hue method allowed participants to read the uncertainty values from the legend and reach their decision more easily. The inclusion of multiple hues was thus welcomed by participants, which supports the statement of Brewer (1994) that a sequential colour scheme can contain more than one hue.

During the post-test participants were also asked to indicate how important the map legend was for their decision-making strategies. Results of this question are presented in Chapter 4.3.4.

4.3. RQ2: Uncertainty Communication

The goal of RQ2 is to investigate the influence of the uncertainty communication method on the damage estimate, the response time as well as the decision-making strategies. To answer RQ2, the data of the two communication groups were compared. Therefore, each group contained results of both uncertainty visualization methods as the communication method is a between-group variable in the study setup. The two samples are thus independent (Field et al., 2012).

4.3.1. Damage Estimate

The overall mean damage estimate was slightly higher for the numerical group (mean: 4.24, SD: 2.10) compared to the verbal group (mean 4.11, SD: 2.08) (see Table 13). Nevertheless, the boxplots in Figure 41 show a rather similar pattern for both communication methods.

Table 13: Results of the damage estimate overall and per communication method.

	Normality	Mean	Median	SD
Overall	p-value < 0.001	4.18	4	2.09
Numerical	p-value < 0.001	4.24	4	2.10
Verbal	p-value < 0.001	4.11	4	2.08

Potential differences in the damage estimates due to the uncertainty communication method were investigated statistically. As the communication group samples were independent and the damage estimates were not normally distributed, a non-parametric Mann-Whitney U Test was appropriate. R studio does not support such a test but the Wilcoxon rank-sum test can be used alternatively (Field et al., 2012).

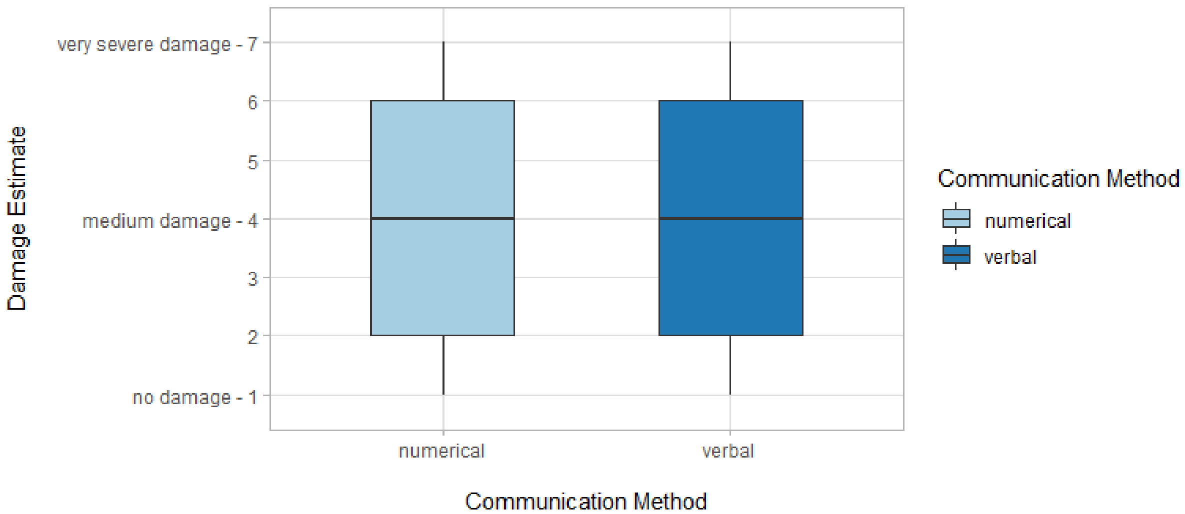


*Figure 41: Boxplot of the damage estimate per communication method. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

Just like the Wilcoxon signed-rank test, H_0 of the Wilcoxon rank-sum test was that the means of two sampled did not differ and the H_A stated that there was a difference in the means. As the p-value of 0.187 resulting from this test is larger than 0.05, H_0 was failed to be rejected and it is assumed that no difference in the damage estimate existed between the two communication methods.

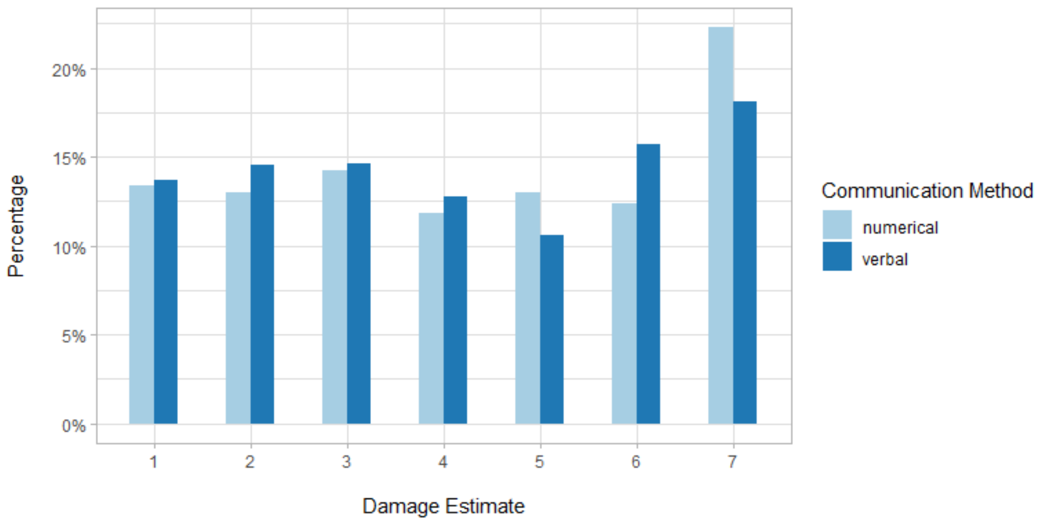


Figure 42: Relative distribution of damage estimate per communication method.

When looking at how damage estimates were relatively distributed in Figure 42, some differences are observable. The highest damage category 7 – *very severe damage* as well as category 5 – *substantial damage* were both chosen more often by the numerical group, while the remaining categories were more frequently chosen by the verbal group.

4.3.2. Response Time

Table 14 shows that participants in the verbal group took longer to make the decision (mean: 14s, SD: 8.34s) compared to the numerical group (mean: 12.32s, SD: 6.91s). This observation can also be made when looking at the boxplot in Figure 43.

Table 14: Results of the response time [s] overall and per communication method.

	Normality	Mean	Median	SD	Min	Max
Overall	p-value < 0.001	13.14	11.43	7.69	1.90	67.13
Numerical	p-value < 0.001	12.32	11.03	6.91	1.90	53.51
Verbal	p-value < 0.001	14.00	11.90	8.34	2.00	67.13

The range of the response times were similar between the two communication groups. The boxplot also shows that the distribution of response times was slightly broader for the verbal groups compared to the numerical one, as the 25th and the 75th percentile are further apart.

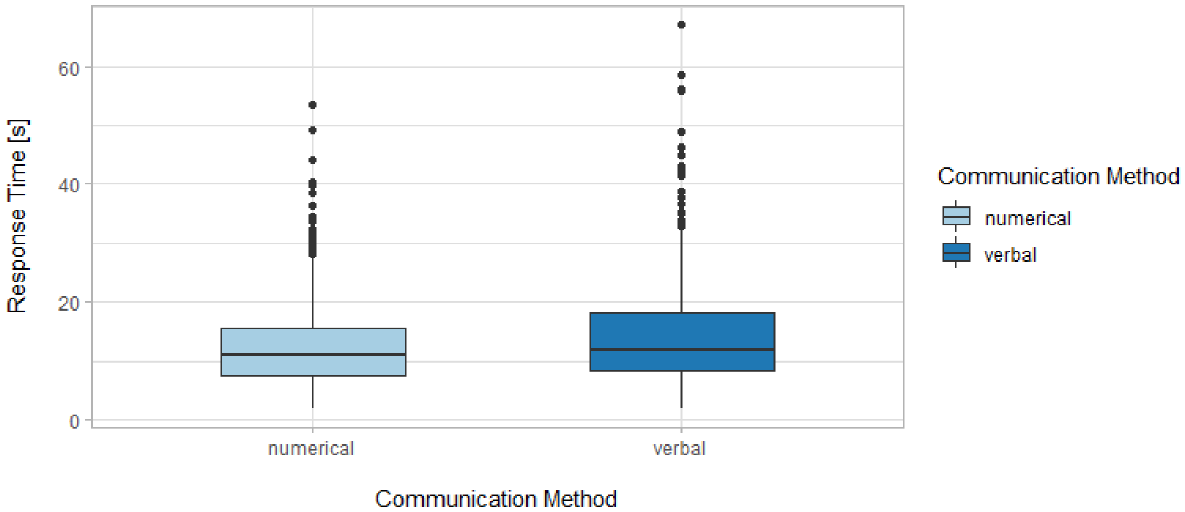


*Figure 43: Boxplot of the response time [s] per communication method. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

The difference in the means was again investigated statistically with a Wilcoxon rank-sum test. The p-value of 0.0001 indicated that the null hypothesis can be rejected and that there was a significant difference in the response time between the communication groups. The analysis of the effect size (r = 0.09), however, showed that the effect was again rather small.

4.3.3. Eye Tracking

The total duration of fixations per AOI in Figure 44 of course shows a similar picture as discussed in the context of RQ1. Therefore, the focus of this chapter will only be on the differences between the two communication methods.

The figure clearly shows that the largest differences can be observed for the AOIs covering the map display, the uncertainty legend and the uncertainty label. The duration of fixation recorded for the numerical groups exceeded the one of the verbal group in these AOIs. This of course reflects the difference in response time just discussed above. The between-group differences of eye tracking metrics per AOI were investigated statistically through independent t-test or Wilcoxon rank-sum tests depending on whether the samples were normally distributed or not. The results in Table 15 show that significant differences in the mean total duration of fixations were found in the following AOIs: Colour scheme, contour label, house location, Likert scale, uncertainty label, uncertainty legend and scale bar. It might seem surprising that the difference for the map display AOI did not yield a significant result. However, this is likely due to the large standard deviations (numerical: 5.05s, verbal: 6.62s) compared to the difference of means (0.85s) in this AOI. Taking into account the effect sizes, it is evident that only the differences in the uncertainty legend and uncertainty label AOI can be considered relevant.

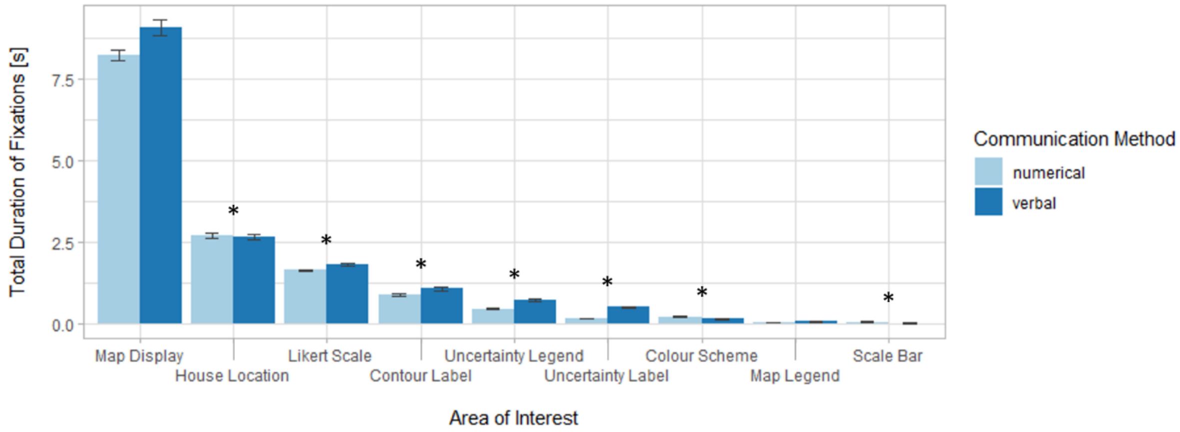


*Figure 44: Mean total duration of fixations [s] per AOI and communication method. Error bars indicate +/- 1 standard error. The * symbol indicates statistically significant differences.*

To create a transparent picture, the total duration of fixations per AOI was again normalized with the AOI area to account for a bias introduced by the different sizes of the AOIs (see Figure 45). Of course, results reflect the same pattern as in Figure 37 with the house location AOI containing the longest total fixation durations. Focusing on the two communication methods it becomes clear that only a slight difference in the time spent in this AOI was observed.

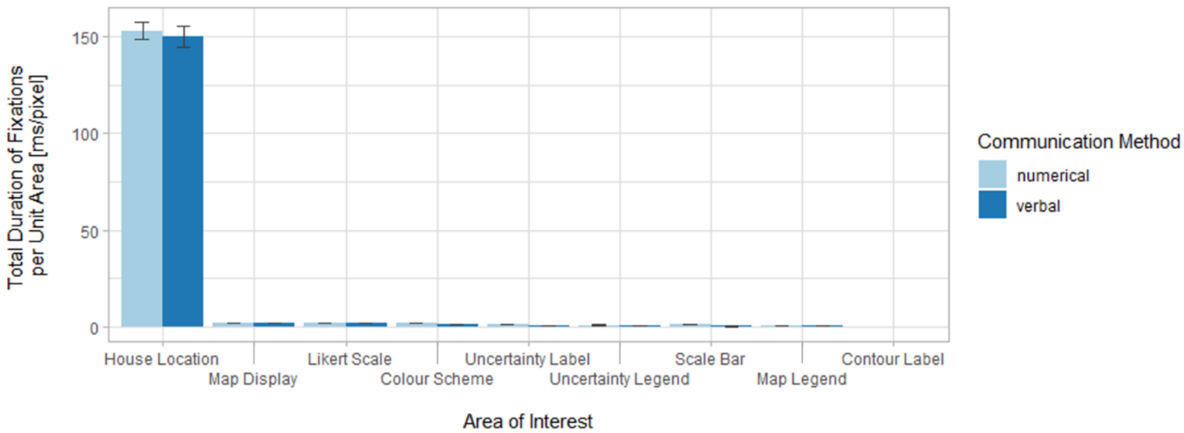


Figure 45: Mean total duration of fixations [ms/pixel] per AOI and communication method, normalized by the AOI area. Error bars indicate +/- 1 standard error.

Table 15: Statistical results of the comparison of means for two eye tracking metrics between the communication methods for each AOI. Italic entries symbolize that the difference was significant and, therefore, the effect size r is reported.

AOI	Total Duration of Fixations	Number of Visits
Colour Scheme	*p-value = 0.008 (r = 0.064)*	p-value = 0.063
Contour Label	*p-value = 0.027 (r = 0.053)*	*p-value = 0.038 (r = 0.050)*
House Location	*p-value = 0.007 (r = 0.065)*	*p-value = 0.027 (r = 0.053)*
Likert Scale	*p-value = 0.011 (r = 0.061)*	p-value = 0.710
Map Display	p-value = 0.209	p-value = 0.806
Map Legend	p-value = 0.085	p-value = 0.086
Uncertainty Label	*p-value < 0.001 (r = 0.268)*	*p-value < 0.001 (r = 0.238)*
Uncertainty Legend	*p-value < 0.001 (r = 0.105)*	*p-value < 0.001 (r = 0.087)*
Scale Bar	*p-value < 0.001 (r = 0.084)*	*p-value < 0.001 (r = 0.083)*

The mean number of visits recorded in the AOIs was rather similar between the numerical (mean: 14.23 visits, SD: 7.93) and the verbal group (mean: 14.98 visits, SD: 8.28). The communication method led to the largest differences in the number of visits for the following AOIs: House location, contour label, uncertainty legend and label (see Figure 46). These AOIs as well as the scale bar were also the ones where significant differences were observed. However, only the effect observed for the uncertainty label can be considered important.

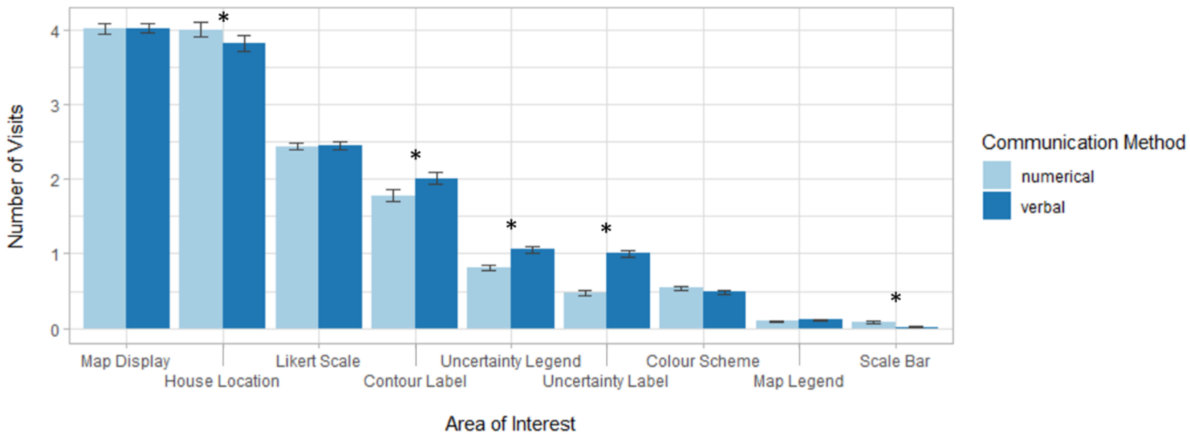


*Figure 46: Mean number of visits per AOI per communication method. Error bars indicate +/- 1 standard error. The * symbol indicates statistically significant differences.*

Comparing the number of visits per Likert scale AOI shows a rather even result between the two communication methods (see Figure 47). Differences worth mentioning were found in level 2 – *very minor damage* and 3 – *minor damage*, which were more often visited by the verbal group, as well as level 5 – *substantial damage* and 7 – *very severe damage*, which were more frequently visited by the numerical group. These differences are in accordance with those in the relative distribution of the damage estimates in Figure 42. Therefore, the differences in the number of visits between Likert scale AOIs reflected the differences in damage estimates for the two communication methods.

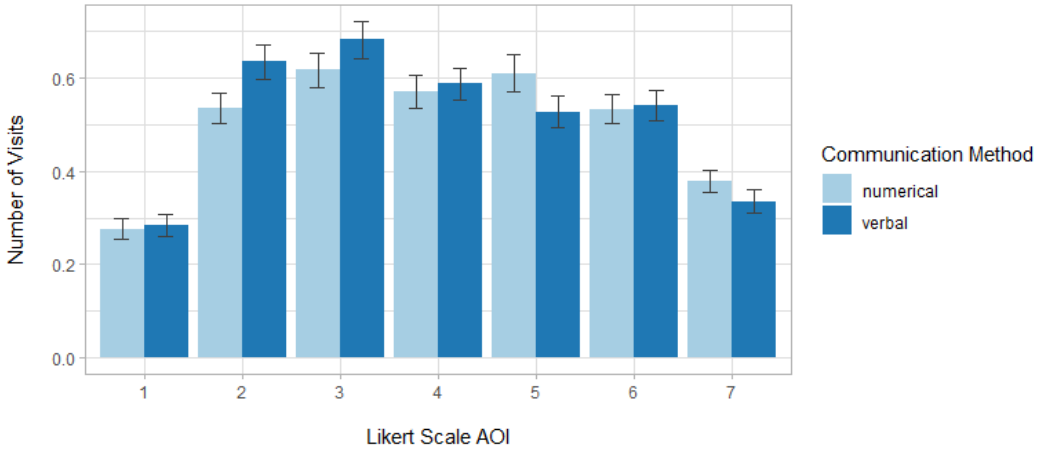


Figure 47: Mean number of visits per Likert scale AOI amd communication method. Error bars indicate +/- 1 standard error.

Similar patterns can also be observed in the qualitative eye tracking data provided through heatmaps in Figure 67 in the Appendix. The relative distribution of fixations clearly cluttered on the map display and the different house locations. The patterns detected in the map displays closely resembled each other between the two groups, which again shows that participants strongly focused on the task relevant house location. The Likert scale pattern just described can also be observed. Additionally, it is evident that the verbal groups spent more time looking at the uncertainty labels compared to the numerical group.

4.3.4. Post-Test

Participants were asked to rate the importance of the map legend to investigate whether there was a difference in how the map legend was perceived between the communication groups. The average importance rating was 4.12 (SD: 1.00) on a scale of 1 – *not important* to 5 – *very important*. These values only varied slightly between the numerical (mean: 4.05, SD: 0.72) and the verbal group (mean: 4.19, SD: 1.25). The relative distribution of the importance ratings is displayed in Figure 48.

A statistical investigation through a Wilcoxon rank-sum test (samples were not normally distributed) resulted in a p-value of 0.18 which shows that there was no significant difference in the scores. It can thus be concluded that no matter if the map legend included numerical or verbal uncertainty labels, it was rated as similarly important by participants. More detailed information on the positive and negative aspects of the map legend mentioned by participants can be found in Table 35 in the Appendix.

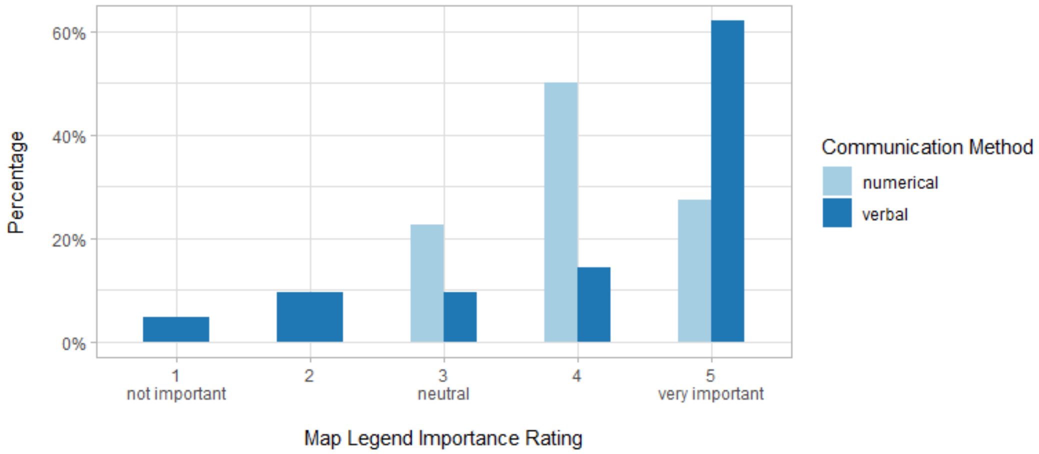


Figure 48: Relative distribution of map legend importance ratings per communication method.

4.4. RQ1 & RQ2: Interaction between Visualization & Communication Method

ANOVA analyses are frequently used to investigate interactions between multiple variables. These analyses require normality of the data and homoscedasticity (homogeneous variances) between the samples (Field et al., 2012). However, these preconditions are frequently violated when collecting data in the form of Likert scale responses in human-computer interaction (HCI) studies, leading to non-parametric data (Wobbrock et al., 2011). As it was shown above, this is exactly the case in this study as the damage estimate and response time data did not follow a normal distribution (homoscedasticity held in all cases except for the response time between communication methods).

There are three options on how to deal with this problem. Firstly, one could simply neglect the fact that ANOVA preconditions are violated and risk an increased Type I error in the test results. Secondly, one could only perform one-way tests which do not take into account potential interactions between variables. Luckily, there is a third option which involves the Aligned Rank Transform tool (ARTool) developed by Wobbrock et al. (2011). This tool performs a data alignment. Through this procedure, regular ANOVA analyses can be performed even if the initial data violates ANOVA assumptions. As the study at hand follows a mixed-design, interactions must be analysed through linear mixed-effects models (LMMs). LMMs investigate how a response variable (e.g., the damage estimate) is related to additional variables (e.g., visualization method or communication method) (Magezi, 2015). This allows to simultaneously analyse results collected through within-group variables (the uncertainty visualization method) and between-group variables (the uncertainty communication method) (Ruginski et al., 2016). As one can imagine, these types of analyses can become complex rather quickly. Therefore, the focus in this chapter is to investigate potential interactions mainly with the help of interaction plots. The results of the LMMs are briefly mentioned and interpreted. More detailed results such as the model outputs and the post-hoc investigation of contrasts can be found in the Appendix R.

4.4.1. Interaction Effects on Damage Estimate

First of all, the combined effect of uncertainty visualization and communication on the damage estimate were investigated for the damage estimate. Table 16 gives an overview of the descriptive statistics for all the different combinations of uncertainty visualization and communication methods. Solely based on these numbers no clear pattern can be observed except that the mean values did not spread widely.

Table 16: Results of the damage estimate across both visualization and communication methods.

	Normality	Mean	Median	SD
Overall	p-value < 0.001	4.18	4	2.09
Numerical & single-hue	p-value < 0.001	4.16	4	2.10
Numerical & multi-hue	p-value < 0.001	4.33	4	2.12
Verbal & single-hue	p-value < 0.001	4.04	4	2.10
Verbal & multi-hue	p-value < 0.001	4.19	4	2.05

A p-value of below 0.05 is necessary to state that the variables significantly affected the fit of the model (Magezi, 2015). Running an ANOVA analysis with the ARTool showed that the interaction between the visualization and the communication method did not significantly influence damage estimates ($F(1,1675) = 0.0046$, $p = 0.946$).

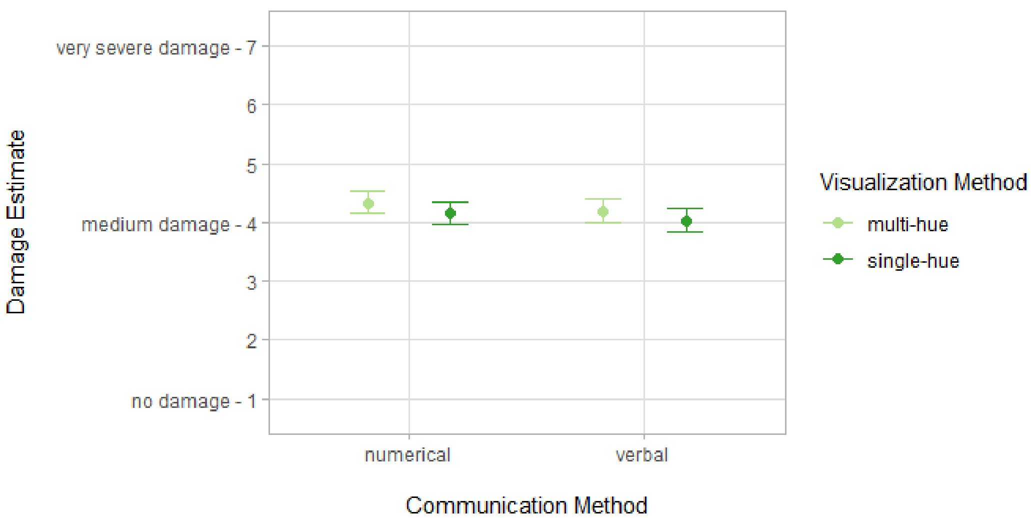


Figure 49: Interaction of communication and visualization method in terms of the damage estimate. Error bars show the lower and upper Gaussian confidence limits.

To interpret the results of the mixed models more easily, interaction plots can be created. They display the effects of different independent variables on the dependent variable under investigation graphically. Figure 49 shows this plot for the interaction discussed statistically above. The plot shows that the multi-hue stimuli resulted in a slightly higher damage estimate compared to verbal method for both communication methods. However, the communication method did not seem to have an influence on the damage estimate within each visualization method. This is of course in accordance with the statistical results presented above.

4.4.2. Interaction Effects on Response Time

The same analysis was conducted for the response time. Again, the descriptive statistics are presented in Table 17 to give an overview. As already discussed above, a higher response time was detected for the verbal group for both visualization methods. Interestingly, mean response times were higher for the single-hue stimuli in the numerical group, while the verbal group took longer to make decisions with the multi-hue stimuli.

Table 17: Results of the response time [s] across both visualization and communication methods.

	Normality	Mean	Median	SD	Min	Max
Overall	p-value < 0.001	13.14	11.43	7.69	1.90	67.13
Numerical & single-hue	p-value < 0.001	12.56	11.16	7.09	1.90	53.51
Numerical & multi-hue	p-value < 0.001	12.08	10.96	6.73	2.02	44.14
Verbal & single-hue	p-value < 0.001	13.50	11.45	8.08	2.00	67.13
Verbal & multi-hue	p-value < 0.001	14.50	12.85	8.57	2.11	56.17

The results of the ARTool mixed model run showed that a significant interaction was found between the communication and the visualization method ($F(1,1675) = 6.86$, $p = 0.009$). Consequently, the effects were analysed in a post-hoc analysis. Post-hoc procedures are used to conduct pairwise comparisons of all levels of variables (e.g., comparing numerical/multi-hue results with numerical/single-hue results). Importantly, the level of significance is corrected for each test through a Bonferroni correction to keep the chance of making a Type I error at 0.05 (confidence level α) (Field et al., 2012). The results obtained in the post-hoc analysis were very surprising as no contrast was found to have a significant effect on the response time (see Table 36 in the Appendix). In this case the interaction plot in Figure 50 can give more insight.

The plot shows that the response time was shorter for the numerical group compared to the verbal one, which is of course in accordance with the results observed above. Therefore, the effect of the communication method on the response time overall was clear. However, the effect of the visualization method on the response time seemed to be inconsistent between the two communication groups. While response time was lower for the multi-hue stimuli in the numerical group, the same stimuli led to a higher response time in the verbal group. This switch in the slope between the two visualization methods was likely the reason why a significant interaction between the communication and visualization method was found but none of the contrasts were found to be significant. Looking at it from the perspective of the visualization method, response time seemed to be more highly influenced by communication method for the multi-hue stimuli compared to the single-hue stimuli.

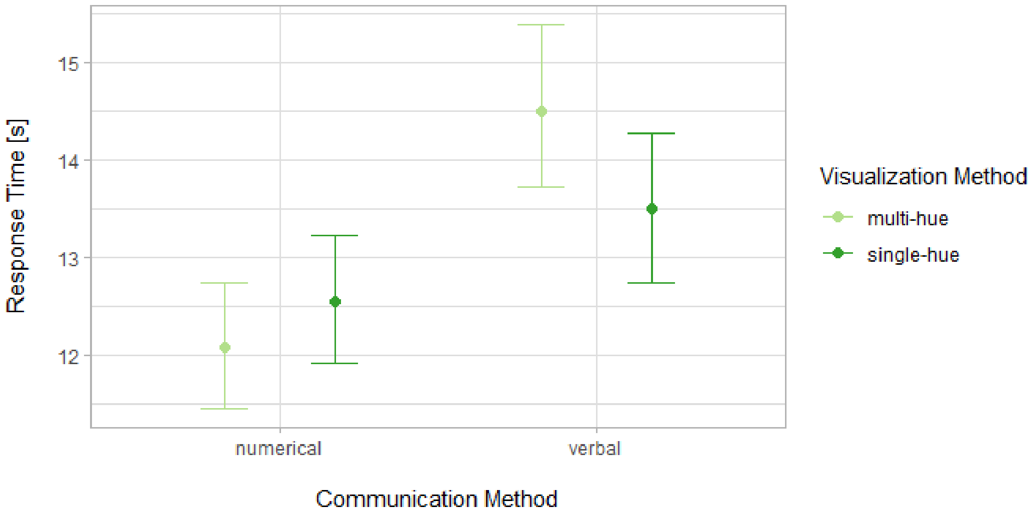


Figure 50: Interaction of communication and visualization method in terms of the response time [s]. Error bars show the lower and upper Gaussian confidence limits.

This overall pattern was also observed when calculating the effect size through the Cohen's d for the communication (d = 0.23) and the visualization method (d = 0.04). This value can be interpreted similarly to the effect size r. Therefore, both variables seemed to have a small effect on response time. However, the effect of communication method was calculated to be larger than the one caused by the visualization method. The low effect of the visualization method can likely be traced back to the inconsistent effect causing the change of slope discussed above.

In conclusion, no interaction was found between the communication and the visualization method for the damage estimate. For the response time, the interaction was significant. Yet, the effect of the visualization method based on the way uncertainty was communicated was inconsistent.

4.5. RQ3: Influence of Additional Information

The aim of the third research question is to investigate whether it is the uncertainty data itself that guides decision-making or if other information has an influence as well. The slope at the house location and the distance between the house and the riverbed are the two parameters for which it was controlled for in this study.

4.5.1. Collinearity among Parameters

Based on the debris flow introduction in Chapter 2.9 as well as the discussion of the data workflow in Chapter 3.4.1.1, it became evident that the three parameters of uncertainty value, slope and distance could be somewhat related. To investigate whether this was the case, a correlation analysis was conducted. The descriptive statistics on these three parameters are displayed in Table 37 in the Appendix R. The Spearman's correlation coefficient r_s, suitable for non-normally distributed samples, was calculated to investigate the potential collinearity of two parameters at the time (Field et al., 2012). The analysis showed that both the pairs uncertainty value / slope (r_s = -0.11) and slope / distance (r_s = -0.26) were weakly negatively correlated. Comparing the damage estimates based on these parameters was thus considered to be possible. However, the uncertainty value and the distance were strongly correlated (r_s = 0.75). This meant that these parameters were collinear and thus analysing the effects of these parameters on the damage estimate in a single linear mixed-effect model would not yield reliable results.

4.5.2. Influence of Parameters on Damage Estimate

The next step included the application of LMMs as described in Chapter 4.4, which enabled the investigation of the effects of different parameters on the damage estimates. Based on the collinearity analysis it was clear that the uncertainty value and the distance must not be included in one model at once or otherwise results could be biased by their collinearity. Additionally, rank deficiency was found in the three-way interaction model, which led to an exponentially growing number of predictors. Therefore, it was also not possible to simultaneously assess the interaction of uncertainty value, slope and distance in a single model. Alternatively, the effects of these parameters on the damage estimate were investigated one at a time as well as pairwise.

Just like the discussion of the LMMs above, the results presented in this chapter will focus on the graphical interpretation of interaction plots and brief explanations of the LMM outputs. Detailed model outputs and results of post-hoc analyses can be found in the Appendix R.

4.5.2.1. Effect of Uncertainty Value on Damage Estimate

Calculating the correlation between the uncertainty value and the damage estimate led to a Spearman's correlation coefficient of r_s = -0.88 (see Table 18). This represents a strong negative relationship: The higher the uncertainty value, the lower the estimated damage in the decision outcomes (see Figure 51). Standard deviations of the damage estimates were similar for the three uncertainty value categories (SD low: 1.10, SD medium: 1.19, SD high: 1.17).

Table 18: Spearman correlation coefficient r_s for the relationship between the uncertainty value and the damage estimate overall and aggregated by the visualization and communication methods.

	Overall	Single-hue	Multi-hue	Numerical	Verbal
r_s	-0.88	-0.87	-0.89	-0.87	-0.88

The LMM output (see Table 38 in the Appendix) showed that uncertainty value significantly influenced damage estimates ($F(2, 1675)= 2210.8$, $p < 0.001$). The effect sizes for all three contrasts (low / high: d = 3.68, medium / high: d = 1.38 and low / medium: d = 2.30; see Table 39 in the Appendix) were large, which means that uncertainty value had a strong effect on the damage estimates.

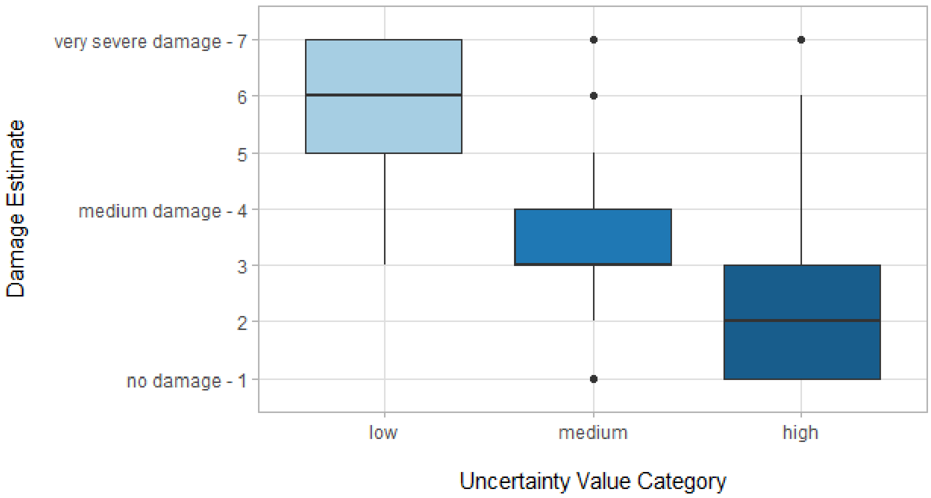


*Figure 51: Boxplot of the damage estimate per uncertainty value category of house locations. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

4.5.2.2. Effect of Slope on Damage Estimate

Investigating the relationship between the slope at the house location and the damage estimate showed that a weak positive correlation existed (r_s = 0.16). This means that a steeper slope led to a slightly higher estimated damage. This relationship can also be observed in Figure 52. It must be noted that the standard deviations of damage estimates were rather high (SD gentle slope: 2.15, SD moderate slope: 1.82, SD steep slope: 2.31), which could explain the weak relationship.

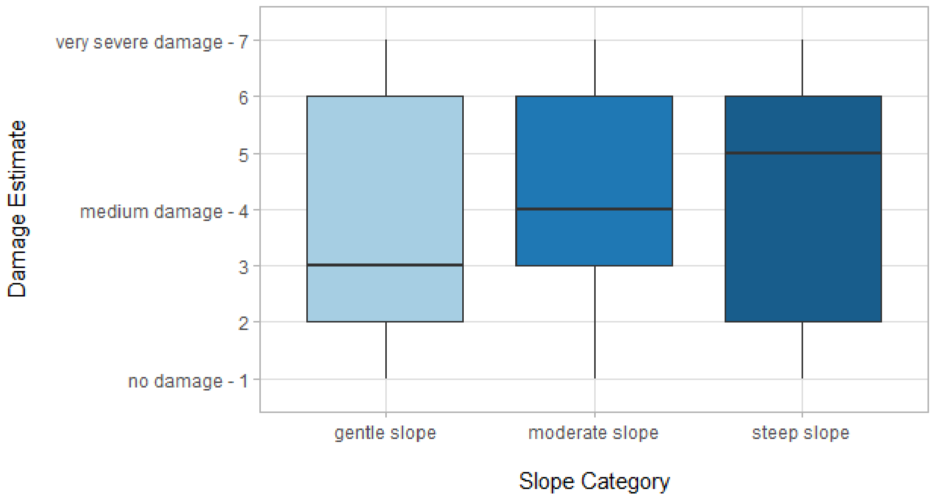


*Figure 52: Boxplot of the damage estimate per slope category of house locations. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

Running the LMM with slope as the only predictor of the damage estimate showed that it had a significant influence ($F(2, 1675) = 14.215$, $p < 0.001$) (see Table 40 in the Appendix). Because this result was significant, the effect sizes of the different contrasts (gentle slope / steep slope, moderate slope / steep slope and gentle slope / moderate slope) were investigated in a post-hoc analysis. The complete results of this analysis can be found in Table 41 in the Appendix. Through this analysis it became clear that the effect of the contrasts gentle slope / steep slope ($d = 0.32$) and gentle slope / moderate slope ($d = 0.23$) are small to medium, while the effect of moving from a moderate to a high slope had a negligible effect on damage estimate.

4.5.2.3. Effect of Distance on Damage Estimate

When investigating the effect of the distance between the house location and the riverbed on the damage estimate, the Spearman's correlation coefficient r_s of -0.72 clearly indicated that a moderate positive relationship existed. The higher the distance between the house and the riverbed, the lower was the estimated damage. Figure 53 shows this very clearly. Compared to the slope, the standard deviations of damage estimates in the different distance categories were smaller (SD short distance: 0.82, SD medium distance: 1.39, SD long distance: 1.48).

Table 42 in the Appendix shows the results of the ARTool model run, which indicates that distance also significantly influenced damage estimate ($F(2, 1675)= 1643.8$, $p < 0.001$). The effect sizes of the three contrasts short / long distance ($d = 3.55$), medium / long distance ($d = 1.48$) and short / medium distance ($d = 2.07$) were all found to be large (see Table 43 in the Appendix). This means that the distance from the riverbed had a strong influence on the damage estimates.

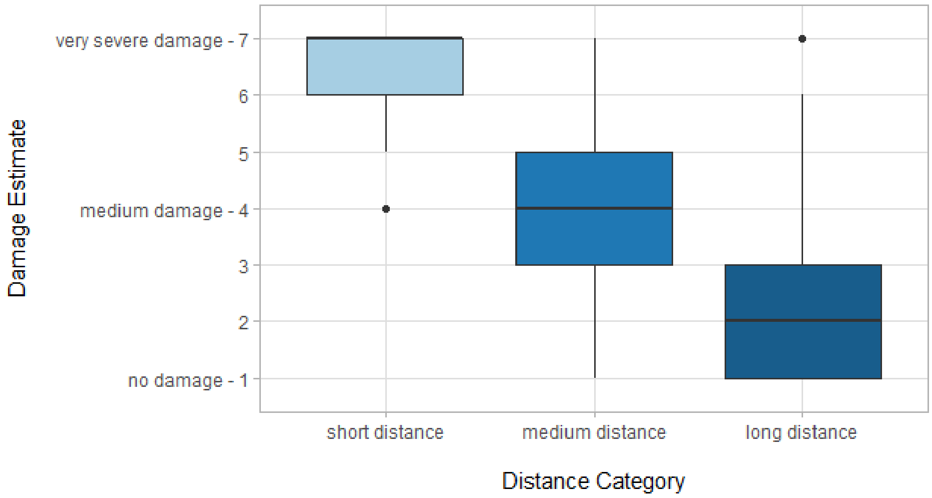


*Figure 53: Boxplot of the damage estimate per distance category of house locations. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

4.5.2.4. Interaction between Slope & Distance

After having investigated the effect of single house location parameters on the damage estimates, their potential interactions were compared pairwise. First, the potential interaction between the slope and the distance category was calculated. The output of the ARTool model can be found in Table 44 in the Appendix and shows that the two parameters significantly interacted ($F(4, 1669) = 154.65$, $p < 0.001$). Again, a post-hoc analysis was conducted to investigate the different contrasts of the model. As the output of the post-hoc analysis is not trivial to read it is included in Table 45 in the Appendix for the sake of completeness. However, the interaction plot in Figure 54 is considered more insightful to interpret the interaction at hand.

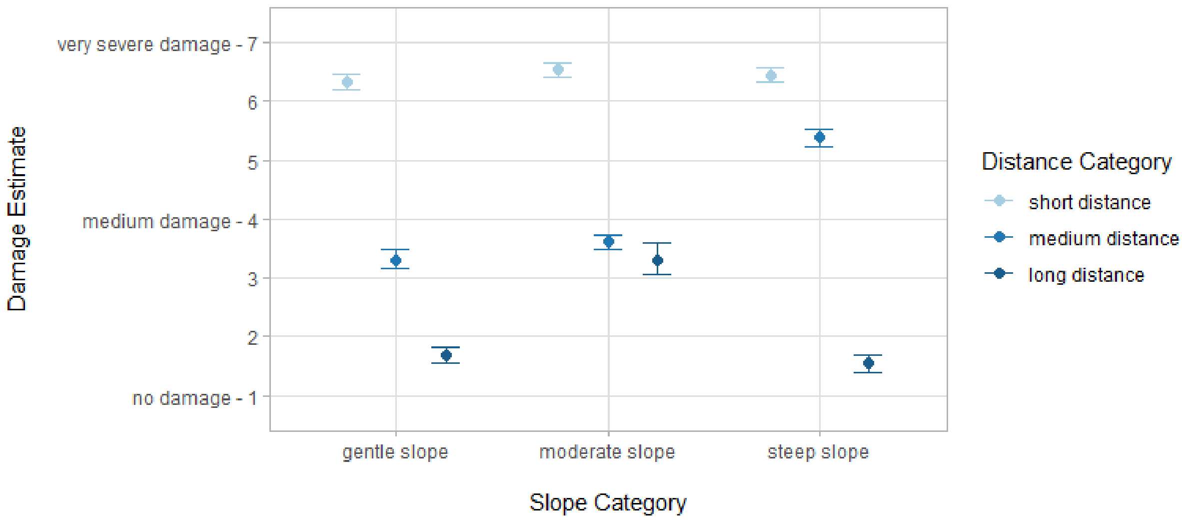


Figure 54: Interaction of slope and distance category in terms of the damage estimate. Error bars show the lower and upper Gaussian confidence limits.

The plot clearly shows that no matter the slope, participants rated the damage similarly high if the house was located close to the riverbed. A steep slope led to lower damage estimates than moderate slopes at long distance locations. It is visible that damage estimates continuously decreased with distance for all slope categories. Interestingly, if slope was moderate damage estimates were quite similar at medium and long distances. Additionally, damage estimates for medium distances were comparatively high at steep slopes. A steep slope also led to a strong increase in damage estimates at medium distances. It can, therefore, be concluded that the effect of distance – an increase in the distance leading to a lower damage estimate – can be observed across all slope categories. The effect of the distance also seemed to overweigh the one caused by slope, which was inconsistent at times. Effect sizes can only be calculated for a single predictor at a time in the ARTool package. The effect sizes of the slope and distance contrasts were already discussed above.

4.5.2.5. Interaction between Slope & Uncertainty Value

When calculating the ARTool model for the potential interaction between the slope and the uncertainty value regarding the damage estimate (see Table 46 in the Appendix), it was found that a significant interaction occurred between the slope and the uncertainty value ($F(3, 1670) = 9.1736$, $p < 0.001$). The significant interaction led to a post-hoc analysis whose results are presented in Table 47 in the Appendix for the sake of completeness. To illustrate the interaction of slope and uncertainty value, the interaction plot in Figure 55 is considered to be more convenient.

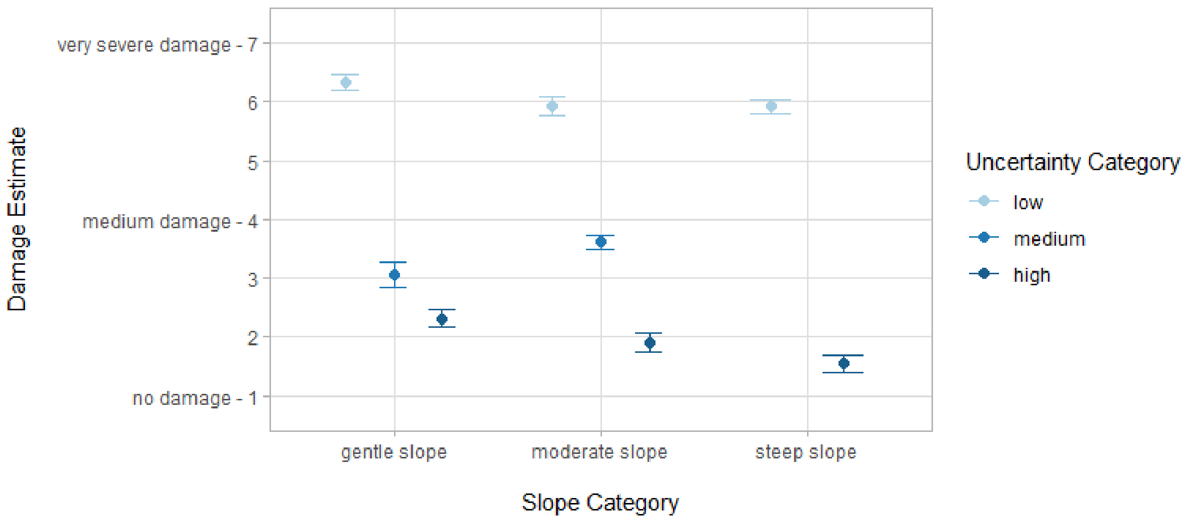


Figure 55: Interaction of slope and uncertainty category in terms of the damage estimate. Error bars show the lower and upper Gaussian confidence limits.

The interaction plot shows that the damage estimate rose as the uncertainty category decreased. Therefore, the more uncertain the value at the house location was, the lower participants judged the damage. If uncertainty was very low, meaning that it was highly likely that the house would be affected, participants chose values around 6 – *severe damage* and 7 – *very severe damage* no matter what the slope was. An increase in slope also led to an increase in damage estimate for the medium uncertainty category. One might notice that no data is plotted at steep slope / medium uncertainty. This can be traced back to the lack of the 20 house locations in this combination of slope and uncertainty value. Moving on to the high uncertainty locations, a decrease in damage estimate with an increasing slope was detected. The conclusion of this interaction analysis is thus that uncertainty seemed to influence the damage estimate more strongly compared to the slope. This is manifested through the continuous increase in damage estimate with decreasing uncertainty value. The pattern caused by the slope category on the other hand was again rather inconsistent. Again, the effect sizes of the slope and uncertainty value contrasts were already elaborated on above.

4.5.3. Post-Test

When asked to rate the importance of the contour lines for their decision-making process on a scale from 1 – *not important at all* to 5 – *very important*, participants gave a mean rating of 4.65 (SD: 0.57). This score was almost equal for the two communication groups (numerical: mean = 4.64 (SD: 0.58), verbal: mean = 4.67 (SD: 0.58)). Consequently, no statistically significant difference was found (Wilcoxon rank-sum test: p-value = 0.844). This means that participants of both communication groups rated the importance of the contour lines similarly high. Figure 68 in the Appendix also shows that no rating below the level of 3 – *neutral* was chosen and that the distribution between the levels of 3, 4 and 5 were very similar for both communication methods.

Table 19: Overview of positive and negative aspects regarding the contour lines in the map displays including the number of mentions overall and per communication group (N = numerical, V = verbal).

Pro	Total	N	V	Contra	Total	N	V
Speed of debris flow	15	7	8	Colour scheme more important	2	1	1
Terrain overview	14	8	6	Contour lines only discovered / taken into account in later trials	2	1	1
Damage estimation	10	6	4				
Direction of debris flow	8	4	4				
Force of debris flow	2	2					
Volume of debris flow	1	1					

An overview of the positive and negative aspects mentioned by participants regarding the contour lines can be seen in Table 19. On the positive side, contour lines helped participants to estimate debris flow parameters such as its speed, direction and the resulting damage as well as getting an overview of the terrain. Two participants respectively mentioned that they put more focus on the colour scheme and that they only took the contours into account in later trials.

4.5.4. Boundary Effect

In the context of the hurricane cone of uncertainty, Padilla et al. (2021a) reported on multiple studies where participants interpreted areas contained by the cone differently from ones outside of it. This boundary effect causes the impressions of a safe zone spanning over areas which are outside the cone of uncertainty. Interestingly, the boundary effect was found to occur despite a detailed explanation of what the cone represents. In Ruginski et al.'s study (2016) investigating different types of visualizations for the hurricane cone of uncertainty, between 28% and 71% of participants (depending on the visualization method) agreed with the following statement: "Areas not shown in blue [the cone] are not predicted to be damaged by the hurricane". This again shows that participants interpreted the boundary as a strict border. However, they also found that the boundary effect was smaller for a fuzzy border visualization compared to a discrete boundary. Similarly, Ash et al. (2014) found significantly higher fear ratings and a higher likelihood to take protective action in a tornado warning scenario when the location was displayed inside the warning zone compared to outside of it. They concluded that the design of an uncertainty visualization can influence the size of the perceived danger zone and must thus be chosen carefully.

To investigate a potential boundary effect in the damage estimates, the data collected with stimuli of high-distance house locations were filtered out. Table 20 clearly shows that the mean damage estimate at locations inside the boundary were higher compared to locations right outside the boundary. Another interesting difference, which can also be observed in Figure 56, concerns the standard deviations. While damage estimates for locations inside the debris flow boundary seemed to vary by around 1.5 scores, the values for locations outside the boundary were distributed far less widely.

Table 20: Descriptive statistics of damage estimates at high-distance house location distinguishing between house locations inside (boundary-in) and outside (boundary-out) the debris flow extent.

	Boundary – in			**Boundary – out**		
	Mean	**Median**	**SD**	**Mean**	**Median**	**SD**
Overall	2.98	2	1.54	1.38	1	0.86
Numerical	3.08	3	1.59	1.326	1	0.74
Verbal	2.89	2	1.48	1.44	1	0.97
Single-hue	2.85	2	1.45	1.39	1	0.94
Multi-hue	3.12	2	1.61	1.37	1	0.77

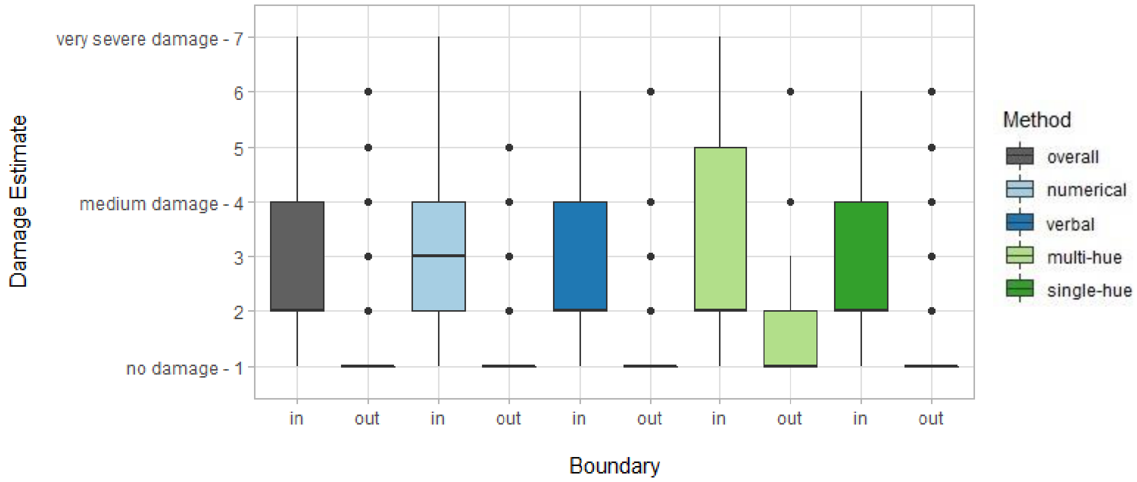


*Figure 56: Boxplot of the damage estimate at high-distance house locations inside and outside the boundary overall and aggregated by visualization and communication methods. The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

Figure 69 in the Appendix also shows that the most frequently chosen damage estimate for locations outside the boundary was 1 – *no damage*, while the damage category 2 – *very minor damage* was most often chosen for high-distance locations inside the boundary.

As all participants solved the task for all the locations inside and outside the boundary the samples are dependent. Since the data was not normally distributed, non-parametric Wilcoxon signed-rank tests were performed to investigate the boundary effect statistically. The mean damage estimates for house locations inside versus outside the boundary were compared for the overall results, each visualization method and each communication method.

Table 21: Results of the Wilcoxon signed-rank tests comparing the damage estimates at high-distance locations inside versus outside the boundary overall and aggregated by visualization and communication methods.

	Visualization Method		**Communication Method**	
	Single-hue	**Multi-hue**	**Numerical**	**Verbal**
p-value	< 0.001	< 0.001	< 0.001	< 0.001
Effect size r	0.51	0.5	0.56	0.50

The Wilcoxon signed-rank test for the overall damage estimates resulted in a significant difference between high-distance locations inside and outside the boundary ($p < 0.001$) with a large effect size of 0.53. The statistical results presented in Table 21 also show that no matter how the data was aggregated, there was always a significant difference between the damage estimates for house locations inside and outside the debris flow boundary. Furthermore, the effect sizes indicate that the significant differences described large effects. It can thus be concluded that the boundary significantly influenced the decision-making of participants leading to significantly lower damage estimates for locations outside the boundary.

4.6. Post-Test

4.6.1. Decision-Making Strategy

At the very beginning of the post-test questionnaire, participants were asked to describe their decision-making strategy in their own words. Additionally, they were asked to indicate which map elements they used when making the decisions. It must be noted that due to the experimental design no specific differences caused by the two visualization methods can be detected as these are a within-group variable. However, differences between the two communication methods as well as trends in the strategies overall can be investigated.

The large majority of participants (n = 40) explicitly mentioned the uncertainty information when describing their decision-making strategy. This means that participants were very aware of the uncertainty information being displayed and took it into account during the decision-making process. Only three participants (one in the numerical and two in the verbal group) did not explicitly mention the uncertainty information. These three participants only based their decision-making on the contours and the distance between the house location and the riverbed. To further structure the decision-making strategies, the following strategy foci were identified: *Contours*, *colour scheme*, *uncertainty label* and *distance*. The author assigned the strategy described by each participant to one of the foci. This classification could be biased due to the manual work involved. An example of one description per strategy focus is provided in Table 23. These examples show that participants included various aspects in their decision-making process. The strategy focus was thus solely determined by the aspect which was perceived to be most important for the description.

Table 22 gives an overview of how the strategy foci were distributed among participants. The table clearly shows that the majority of participants put the focus of their strategy either on the contour lines or on the colour scheme. Interestingly, the distribution of the two communication groups were extremely similar. This indicates that the uncertainty communication method did not clearly affect the decision-making strategy.

Table 22: Overview of the number of participants per strategy focus overall and per communication group.

Strategy Focus	Overall	Numerical	Verbal
Contours	19	9	10
Colour scheme	18	10	8
Uncertainty label	4	2	2
Distance	2	1	1

Table 23: Examples of decision-making strategies with a different strategy focus. The statements were translated to English by the author.

Strategy Focus	Strategy Description
Contours	I first looked at the contour lines to investigate the topography around the house (steep or flat terrain). If contours were narrow, I generally estimated the potential danger to be higher, because I assumed that the speed would be higher compared to a flattening terrain. In a next step, I looked at the uncertainties to estimate how likely it is that an event will affect the area. Based on these two aspects I arrived at my estimation.
Colour Scheme	I first investigated where the top and the bottom [of the terrain] were, then I searched for the darkest shade of green. I then checked whether the house was located in the dark green area (I checked the colour around the house). If the house was in the dark-green area and within the main channel of the debris flow, I assumed that it would be level 6 or 7. If the house was in a brighter area where the debris flow expanded and slope flattened, I gave it a 4. I paid attention to whether the house was protected by hills in the landscape or not. After a while I also started paying attention to the altitude difference the debris flow travelled to estimate the force with which the debris flow arrives at the house.
Distance	It was a mixture of intuition and imagination of the debris flow event. I paid more attention to the location of the house in space as opposed to the uncertainty to estimate the damage. A house which was location right within the debris flow channel would likely be destroyed completely. With the houses at the boundary, I was unsure which damage would be expected.
Uncertainty Label	A) First, I looked at the approximate value on the uncertainty scale. I then quickly checked the contours to get an overview of the map. Then I checked whether the house was at a steep or flat location. To make the decision, I mentally projected the uncertainty scale onto the damage scale, which means that a small uncertainty value leads to a large value on the damage scale. Depending on the topographic location of the house (steep vs. flat, and the width of the debris flow and the uncertainty in flat areas) I then rounded the damage value upwards or downwards. For instance, if the house was in a flat area, I rounded downwards. B) I focussed on the visible changes of colour on the map. If few changes in colour were visible between the upper area (uncertainty = 0%) and a certain point, I expected a (very) severe damage. Sometimes I also focused on the contours to see whether I can imagine the trajectory of the debris flow. If the location was in an area of 50% uncertainty or lower, I decided on the damage level 4 and lower. If the location was outside the boundary, I chose level 1 (no damage).

A similar picture can be seen when analysing the first strategy element mentioned in the descriptions. Again, the colour scheme and the contour lines dominated the picture. Table 24 shows that the numerical group first mentioned the colour scheme most often, while the verbal group first mentioned the contour lines.

Table 24: Overview of the number of participants who first mentioned a strategy element during the decision-making strategy description overall and per communication group.

Strategy Element	**Overall**	**Numerical**	**Verbal**
Colour scheme	16	9	7
Contours	14	5	9
Contour label	8	5	3
Distance	4	2	2
Uncertainty label	1	1	-

One aspect which was evident in most strategy descriptions was the weighing of different aspects. For instance, some participants focused their strategy on estimating the damage through the steepness of the terrain represented through the contours or checked the pseudo-colouring at the house location. In a next step they used another aspect of the map display, such as the distance between the riverbed and the house location, to weigh the damage estimate for locations in a similarly steep or uncertain area. One participant described their weighing of colour scheme and distance as follows: *"First I always checked how dark it [the colour] was around the house and then checked the direction of the debris flow and whether the house was more towards the boundary or in the middle of the potential debris flow. If the house was surrounded by a dark colour, I tended to rank it as more endangered. This was amplified if the house was in the middle of the debris flow and weakened if it was at the boundary"* (translation by the author).

A total of three participants (two from the numerical and one from the verbal group) explained that they made use of threshold to estimate the damage. One participant of the numerical stated that they chose a (very) severe damage level if uncertainty was 0%. If the uncertainty was below 50% they chose a damage level of 4 or below and if the house location was outside the debris flow shape they chose the damage level 1. The other two participants created the thresholds based on the colour schemes. A specific strategy related to these thresholds was mentioned by eight participants. This strategy included an active implementation of the boundary effect described above. These participants actively chose a damage level of 1 – *no damage* if the house was located outside the debris flow shape while choosing a higher damage level if the house was just within the shape.

A few participants reported on a change in their strategy over the course of the study. It was for instance stated multiple times that the contour lines were first neglected and only taken into account in later trials. Additionally, three participants explicitly mentioned that they based their damage judgements of later trials on previous decisions. This means that participants were able to adapt their decision-making strategy over the course of the study and learned from previous decisions.

As mentioned above, differences caused by the two visualization methods could not be systematically assessed. However, one participant explicitly mentioned that they generally judged damages higher for the single-hue colour scheme compared to the multi-hue one. For the multi-hue colour scheme, they chose a larger variety of damage levels.

The map elements which were used by participants to make their decisions were analysed in two ways. Firstly, the strategy descriptions were scanned for the map elements and an overview of the map elements which were explicitly mentioned was created (see Table 25). Secondly, the participants were explicitly asked to indicate whether the used a specific map element from the following list: *Contour lines*, *colour scheme*, *hill shade*, *house location*, *map legend* and *scale bar*.

The first part of the analysis showed that both the contours and the house location were directly mentioned by almost all of the participants. This did not come as a surprise for the house location as this element was of course extremely relevant for the task. However, the number of participants mentioning the contours was certainly higher than expected. Two other map elements which were mentioned very often were the colour scheme and the uncertainty label, which were again task-relevant. The contour labels were mostly used in the context of getting an overview of the map display and investigating the flow direction of the event. The differences between the two communication methods were again minor. The only difference worth mentioning was that numerical participants were more likely to mention the distance between the house location and the riverbed compared to the verbal group.

Table 25: Overview of the number of participants who explicitly mentioned certain map elements overall and per communication group.

Map Element	Overall	Numerical	Verbal
Contours	38	20	18
House Location	37	19	18
Colour Scheme	30	16	14
Uncertainty Label	22	11	11
Contour Label	15	7	8
Distance	14	10	4
Map Scale	2	1	1

To quantitatively investigate the complexity of decision-making strategies, the mean number of map elements mentioned in the strategy description was investigated. This count ranged from 1 to 6 map elements and can be traced back to the different levels of detail with which participants described their strategies. Participants mentioned 3.67 (SD: 1.13) map elements on average. This value was slightly higher for the numerical group (mean: 3.82, SD: 1.26) compared to the verbal group (mean: 3.52, SD: 0.98). Therefore, one could say that the numerical groups took slightly more map elements into account during their decision-making process. This statement is of course only true under the assumption that participants were able to accurately describe their decision-making strategy.

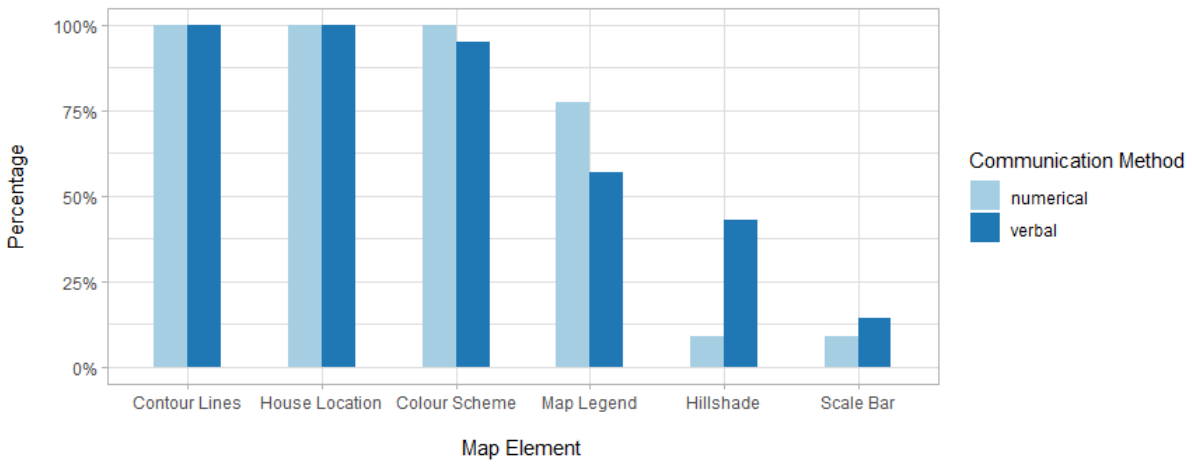


Figure 57: Relative distribution of map elements used during the decision-making process per communication method.

The results of the second part of investigating which map elements were used during decision-making are presented in Figure 57. All participants indicated that they used the contour lines and the house position when making their decision. This is in accordance with the results presented above. Almost all of the participants took the colour scheme into account. Interestingly, the numerical group indicated that they used the map legend more often, while the hill shade was used more often in the verbal group. The latter came as a surprise as none of the participants explicitly mentioned the hill shade or the base map in their decision-making description. Almost none of the participants of both communication groups actively used the scale bar.

4.6.2. Trust Rating

Jenkins et al. (2017) argued that, because natural hazards contain high uncertainty by nature, communicating uncertainty numerically could be seen as too detailed and lead to a decrease in the credibility of communicators. To investigate this, participants were asked to judge how much trust they put into the uncertainty data displayed in the debris flow maps on a scale from 1 – *very low trust* to 5 – *very high trust.*

The relative distribution of trust ratings can be found in Figure 70 in the Appendix. Overall, the trust rating reached an average of 4.21 (SD: 0.91), which is clearly very high. When comparing the two communication groups, it was found that the numerical group rated their trust higher (mean: 4.41, SD: 0.59) compared to the verbal group (mean: 4, SD: 1.14). Although a Wilcoxon rank-sum test showed that the trust ratings between the groups did not differ significantly (p-value = 0.393), it can be concluded that they do slightly vary depending on the communication method.

4.6.3. Task Difficulties

Other authors investigating different ways of communicating uncertainty (e.g., Budescu et al. (2009)) have argued that verbal statements are more vague. This could mean that participants in the verbal group struggled more to make their decision. To get insight into this, participants were asked to rank the task difficulty on a scale from 1 – *not difficult* to 5 – *very difficult* with a neutral position at level 3.

Table 26: Task difficulties mentioned by participants and their respective count.

Difficulty Description	Count
Doubt: doubt in one's decision-making, strategy, consistency and correctness of answer	7
Damage estimation: judging the damage, relating the uncertainty value to a damage level	6
Experience: missing experience regarding the process	3
Map reading: low experience in map reading, difficulty interpreting contour lines	3
Weighing: weighing of different factors (e.g., uncertainty, colour scale, contours)	3
Boundary: judging damages of house locations close to the boundary	2
Interpretation of uncertainty	2
Landscape: missing knowledge on landscape, e.g., vegetation or rocks	2
Uncertainty definition	2
Colour value: differentiating between different values of green	1
Procedure: sitting still	1
No difficulty mentioned	10

The overall mean task difficulty rating was 2.53 (SD: 1.01), which represents a neutral position with a tendency towards *not difficult*. The numerical group gave slightly lower ratings (mean: 2.36, SD: 1.00) compared to the verbal group (mean: 2.71, SD: 1.01). The Wilcoxon rank-sum test investigating this difference statistically yielded a non-significant result (p-value = 0.241). An interesting pattern is visible in Figure 71 (see Appendix R). Only around 10% of the numerical group gave ratings above the neutral category (level 4 or 5), while almost 30% of the verbal participants chose the rating 4, representing a tendency towards judging the task to be difficult. It can thus be concluded that even though the mean rating showed that participants interpreted the task to be neutral or not as difficult, the relative distribution indicated that participants in the verbal group were more likely to judge the task to be difficult. This is in line with the statement regarding the vagueness of verbal uncertainty expressions above.

The difficulties mentioned most often were doubt in their own decision-making strategy and a struggle to relate the uncertainty value to a damage level to estimate the damage. Other difficulties included missing experience, low map reading skills and the weighing of different factors (for a full list see Table 26).

4.6.4. Uncertainty Depiction

Towards the end of the post-test questionnaire, participants were asked to state whether they thought that the depiction of the uncertainty information supported their decision-making or not. The relative distribution of the answers to this question shows that only around 5% of each communication group gave a score below the neutral level of 3 (see Figure 58). They gave their answer in the form of a score from 1 – *impede decision-making* to 5 – *support decision-making*. The overall mean score was 4.42 (SD: 0.91) with very similar results for the numerical (mean: 4.36, SD: 0.95) and the verbal group (mean: 4.48, SD: 0.87). The difference between the two communication groups was investigated with a Wilcoxon rank-sum test and found to be insignificant (p-value = 0.567). This means that the participants of both groups thought that the uncertainty information was beneficial for their decision-making.

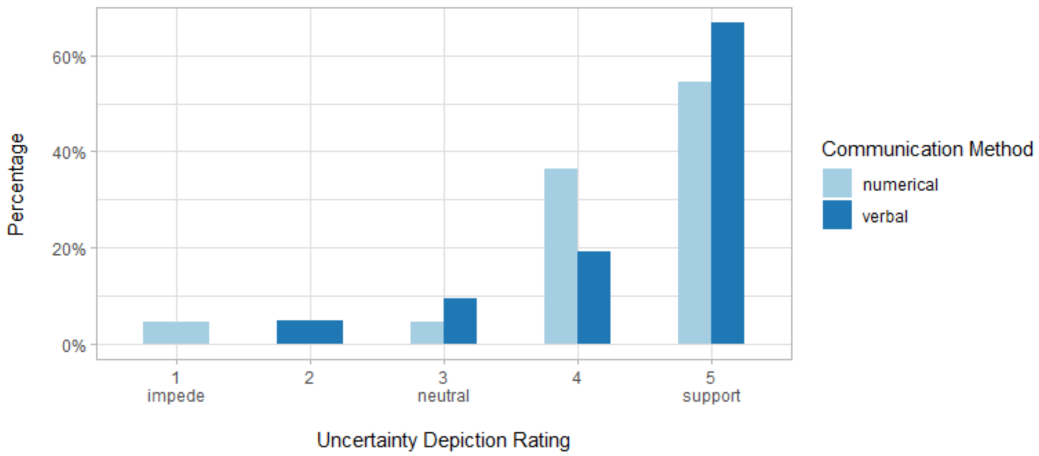


Figure 58: Relative distribution of ratings on the depiction of uncertainty per communication method.

Participants were then asked to elaborate on their decision. 16 participants mentioned that the uncertainty information helped them to distinguish different levels of danger or damage. On the negative side, two participants put forward that they struggled to translate the uncertainty information to the damage scale. Single participants each mentioned that they either did not clearly understand how the uncertainty was defined, that they had trouble interpreting certain levels of uncertainty and that the uncertainty information distracted them from other map content (see Table 27).

Table 27: Overview of positive and negative aspects mentioned by participants regarding the depiction of uncertainty information including the number of mentions overall and per communication group (N = numerical, V = verbal).

Pro	**Total**	**N**	**V**
Distinction of danger / damage	16	7	9
Judge likeliness of the house being reached by debris flow	7	2	5
Indication of affected area	4	2	2
Estimation of debris flow volume	3	2	1
No other previous knowledge of the area	2	1	1
Map display easier to understand compared to only uncertainty numbers	1	1	-
Additional information to support decision making	1	1	-
Prevented overestimation of damage	1	-	1

Contra	**Total**	**N**	**V**
Difficulty to translate uncertainty to a damage level	2	2	-
Uncertainty labels not taken into account	2	-	2
Unclear definition of uncertainty led to confusion	1	1	-
Distracts from other aspects	1	1	-
Difficult to judge high uncertainty	1	-	1
Difficult to judge medium uncertainty	1	-	1

5. Discussion

5.1. General Discussion

5.1.1. Interpretation of the Uncertainty Information

Likely one of the most important observations of this study lies in the way participants interpreted the uncertainty information. As shown above, the uncertainty value strongly correlated with the damage estimations. Therefore, participants were essentially scaling the uncertainty value onto the damage scale. This is a faulty strategy if one strictly considers what the spatial uncertainty on the maps actually meant. The information shows how uncertain it is that a specific location will be affected by a debris flow. Severe damages could occur even if uncertainty in the debris flow modelling is high. Similar results were found in a study by Tak et al. (2014). They observed that participants constructed their own model of uncertainty regarding the boundary between earth layers. They found that this model resembled a normal distribution. This model was then used by participants to make their judgments. Adapting these findings to the study at hand, the model would represent a negative linear association between the uncertainty and the damage level. Probability matching is a concept related to these observations. The idea is that "recipients of the alerts match their responses to an alert with the likelihood of the event occurrence" (Miran et al., 2019). Thus, when a house location was more likely to be affected by a debris flow, participants responded with higher damage estimations.

Interestingly, only two participants reported challenges with interpreting the uncertainty information or its definition. It can, therefore, be assumed that all other participants thought they had understood the uncertainty information correctly. This is in line with the *hard-easy-effect* which states that humans are overconfident when solving a complicated task (Hullman et al., 2019). Additionally, the self-assessments of the post-test questionnaire showed that participants were under the impression that the depiction of uncertainty supported their decision-making. This clearly stands in conflict with the findings just discussed and uncovers that what humans think they do, is not actually what they do.

The pre-test showed that participants had restricted experience in dealing with data uncertainty (mean rating: 2.78, SD: 1.25 on a scale from 1 – *low experience* to 5 – *high experience*). It can thus be assumed that this lack of knowledge caused the misinterpretation of the uncertainty information. As others such as Doyle et al. (2014) and Loucks et al. (2005) have argued, training on how to interpret uncertainty information and how to incorporate it into the decision-making process is crucial to enable truly informed decisions. Publishing more visualizations including uncertainty information is generally encouraged by this author. However, it is equally important that these visualizations are accompanied with formal explanations and instructions for their readers. A very recent study delivered clear support of the positive influence of training on decision-making with uncertainty. Kusumastuti et al. (2022) conducted two experiments to investigate the effect of practice on the accuracy of decision-making in an abstract task including spatial uncertainty. Both experiments showed that decision accuracy increased and response time decreased after subjects went through practice and received direct feedback. Linking back to the decision-making process with uncertainty as introduced in Chapter 2.5, they concluded that the inference step became more successful through the training. These effects were consistent across three different types of visualization (ellipse, interlace and scatter). Additionally, they found that the training on a specific type of uncertainty visualization also had positive effects on decision-making with visualizations which were not part of the practice. This means that training on interpreting uncertainty information with a specific visualization can even have positive effects on peoples understanding of uncertainty overall.

In conclusion, the observations of this study indicate that the non-expert audience had trouble interpreting uncertainty information in a map-based setting. This is in accordance with doubts on visualizing uncertainty found by Fischhoff (2012) and Hullman (2020). Joslyn and Savelli (2021) have argued that humans do not understand the theory behind uncertainty, but that they can use uncertainty information in practice. Yet, the results at hand showed that participants failed to correctly incorporate uncertainty into their decisions. This gives insight into a lack of scientific literacy among non-experts, as introduced in Chapter 3.3 (McMahon et al., 2015). Additionally, the statement expressed by Padilla et al. (2021b), that non-experts have trouble interpreting common formats of uncertainty expressions, is reinforced. It is also true that participants tried to handle the challenging decision-making task by employing heuristics as proposed by Kahnemann and Tversky (1982) (discussed in Chapter 5.1.3). Thus, the task, which was expected to be solved with System 2 processing, as proposed by the dual-process theory in Chapter 2.5, was performed with System 1 processes with the help of heuristics. Despite the misinterpretation of the uncertainty information, the confrontation with uncertainty visualizations certainly contributed to raising the audience's awareness of uncertainty in natural hazards.

5.1.2. Eye Movement Patterns

The main assumption when using eye tracking is that humans process the object they are currently viewing (Just and Carpenter, 1976). Additionally, it is argued that humans attribute most attention to objects which are either visually salient (bottom-up perspective) or which are relevant to the task at hand (top-down perspective) (Itti et al., 1998; Padilla et al., 2018).

The overall results obtained by measuring the eye movements of participants showed that the most time and thus cognitive effort was invested to process the map display (see Figure 36 and

Figure 44). An eye tracking study by Brus et al. (2012) showed that the legend explaining the uncertainty was highly important. This was not the case in this study. Based on the two metrics – total fixation duration and number of visits – the most task-relevant map elements were the map display, including the house location, and the Likert damage scale. The perceived importance of the map display was clearly influenced by its mere size. The normalization of the total fixation duration with the AOI area then showed that fixation duration highly concentrated on the house location AOI. Combining the findings of the raw distribution of total fixation duration and the one normalized by area shows that the house location was the most task-relevant map element. This is in accordance with the design of the task. Surprisingly, the map legends explaining the colour schemes and the associated uncertainty information were only of secondary importance.

The analysis of the number of visits and its comparison with the relative distribution of the chosen damage estimates showed that more visits were registered in the AOIs of the Likert levels 2 – *very minor damage* to 6 – *severe damage* compared to how often these levels were chosen as an answer. The fact that participants were moving their gaze between these levels could be an indicator of more complicated decisions or indecisiveness. The number of visits registered on the Likert levels 1 – *no damage* and 7 – *very severe damage* were surprisingly low compared to the relative distribution of the answers. From these observations it can be concluded that participants did not have trouble making their decision for the highest and lowest damage level. However, when they chose a damage level in between the extremes, they seemed to struggle more and moved their gaze between the levels of the Likert scale.

Comparing the heatmaps of the eye movements (Figure 66 and Figure 67) with the results of the saliency analysis (Figure 26) shows little agreement between them. This can be traced back to the mechanism of the Saliency Toolbox which strictly focuses on the bottom-up processes of scene viewing (Walther and Koch, 2006). Saliency models are often used in eye movement prediction for free viewing studies (Koehler et al., 2014). However, clear limitations of using saliency models to predict eye movement become evident when a specific task is introduced. Through the comparison of actual eye movements with the saliency analysis, it was found that top-down processes introduced by the decision-making task were more important compared to the visual saliency of the map displays (bottom-up processes) in the study at hand.

It is clear that some of these insights would not have been gained solely based on the damage estimates and the response times. Due to its objective nature (Li et al., 2010), the method of eye tracking served as a complementary source of data on the cognitive processes and complemented the self-reports collected in the post-test. Therefore, the value of applying eye tracking technology in map-based decision-making studies was reinforced.

5.1.3. Decision-Making Strategies

As mentioned by Korporaal et al. (2020), studies should collect multiple sources of data to investigate the process of decision-making with uncertainty. The collection of decision-making strategy descriptions is considered to add value to the discussion, especially for the study at hand as the decision outcomes did not correspond with the expectations. One assumption when investigating decision-making strategies is that participants are capable of reflecting on their own behaviour and truthfully reporting on it (Martin, 2008).

The first observation regarding the decision-making strategies was the variety of strategy foci found among the descriptions. It was expected that the uncertainty information represented through the colour scheme and the uncertainty labels would be the most important aspect of the strategies. However, it was found that an almost equal number of participants were attributed to the strategy foci *contours* and *colour scheme*. Only four participants (two of each communication group) were attributed to the strategy focus *uncertainty label*. The high number of participants making use of the contours could be traced back to the high level of map reading skills found in the self-assessment of participants (mean rating: 3.95, SD: 0.87 on a scale from 1 – *low experience* to 5 – *high experience*). The usage of different strategy foci shows that there was no clear way to go about the decision-making task. This reflects the complexity of decision-making with uncertainty information.

The level of detail with which participants described their decision-making strategy was highly variable. While some descriptions were only composed of a few bullet points or a single sentence, others took the form of long paragraphs. This could of course reflect differences in how much cognitive effort was put into the decision and could be a reason for the rather high range of response times.

The map elements which were more often mentioned in the strategy descriptions were the contours, the house location, the colour scheme and the uncertainty label in the order of their mention. It was surprising to see that participants used the contours this often. However, the remaining pattern certainly reflected the task relevance of the different map elements. Nevertheless, the importance of the uncertainty label was expected to be higher. The analysis of the used map elements further showed that the hill shade and the map scale were the least often used for decision-making. This trend was also observed in the strategy descriptions, where they were hardly mentioned. A potential reason for this is that participants were not aware of the hill shade until they were asked to explicitly judge whether they used it or not. The fact that the map scale was hardly mentioned during the strategy descriptions as well as when explicitly asked to indicate the used map element leads to the conclusion that the map scale could be excluded in future map displays. This is rather surprising as quite a few participants mentioned that they used the distance between the house location and the riverbed in their strategies, which is clearly tied to the map scale. However, this shows that the relative distance was more important than the absolute one.

Interestingly, multiple participants reported on a learning effect over the course of the trial. The task became increasingly easy for participants to solve. Independent of whether participants interpreted the uncertainty correctly or not, it is expected that the more experience in dealing with uncertainty information is acquired, the more successful will the decision-making be, especially for a non-expert audience.

As mentioned above, the experimental design prevented an in-depth investigation of differences in decision-making strategy caused by the two visualization methods. Only one participant mentioned that they made their decision slightly different for the two colour schemes. This means that participants either did not notice that the colour scheme changed throughout the study (as explicitly mentioned by two participants) or they applied the same strategy. The latter hypothesis could also explain the small effect of the visualization method on the damage estimates. The eye tracking data showed that the colour scheme was hardly looked at in the map legend. However, 30 out of 43 participants specifically mentioned the colour scheme in their decision-making strategy. Therefore, the participants likely processed the colour scheme directly on the map display instead of using the legend.

No clear pattern was observed regarding the influence of the two uncertainty communication methods – numerical and verbal. Neither the strategy focus nor the mentioned map elements seemed to be affected by the way uncertainty was communicated in the legend. This could be traced back to the lack of eye movements registered on the map legend. Against the expectations, the verbal group did not show a tendency to use the additional information (contours & distance) more heavily in their strategies to compensate for the vague uncertainty expressions.

Even though neither the uncertainty visualization nor the communication methods seemed to have an influence on the decision-making strategies, some interesting findings on the strategies overall must be discussed. Decision-making theories imply that humans use a variety of heuristics when making decisions with uncertainty. The purpose of these heuristics is to simplify the decision-making process (Padilla et al., 2021a). The three most important ones of them were presented in Chapter 2.5. One of these, which could have had an influence on the results, is the availability heuristic (Tversky and Kahnemann, 1974). This would mean that participants who had already experienced a debris flow would likely overestimate the damages. The pre-test showed that only one participant had been personally affected by an event, as their holiday house was damaged. They explicitly mentioned a form of an availability heuristic: *"I basically attributed houses located in the dark green areas to the highest damage level. This could be cause our own* *house was flooded last summer and I know how little it takes to damage a whole floor of a house severely. This is why I felt a lot of sympathy for the people in the house"* (translation by the author). Since only one participant had previously experienced such damages, it can be assumed that the availability heuristic did not have an influence on the results overall. Nevertheless, one aspect which could have influenced participants subconsciously was that extraordinarily high damages caused by various natural hazards were recorded during the year of 2021. The financial costs resulting from these damages were more than ten times higher compared to the previous year and reached around 450 Mio. CHF (Liechti et al., 2022). Clear indications of the representativeness heuristic were not found in the decision-making strategies. However, the subjective expectations of damages caused by debris flows could have had an influence. Thus, participants who interpreted debris flows as a very threatening natural hazard likely judged damages higher compared to ones who did not see debris flows as a threat. Lastly, three participants mentioned that they based their damage estimations on previous trials. This could be seen as an expression of the anchoring heuristic. They used their previous estimations as an anchor and made their decision based on it. This could influence the overall pattern of damage estimates for these participants.

Another shortcut rooted in the decision-making strategies, which relates back to the misinterpretation of the uncertainty information, is the heuristic of attribute substitution. This means that the complicated uncertainty information is mentally swapped out by a simpler representation, which is likely not as accurate (Joslyn and LeClerc, 2013). In its most extreme form, a *deterministic construal error* can occur. In this scenario the uncertainty information is completely neglected and the data is treated deterministically (Padilla et al., 2021a). During the study at hand, participants might have substituted the uncertainty information with the severity of a damage. This could explain why they matched low uncertainties to high damages and high uncertainties to low damages. Similar results were found by Kübler et al. (2020). Their participants interpreted the depiction of uncertainty as indicating different levels of danger instead of the uncertainty regarding the danger zone (more on this in Chapter 5.2.1).

Participants also simplified the decision-making process was through a distance heuristic. This led to the rather strong negative correlation between the distance and the damage estimate in Figure 53. This heuristic represents one of the three geospatial framing effects as discussed by Klockow-McClain et al. (2020). They found that participants interpreted numerical percentages differently, depending on how far away a specific uncertainty level was from the centre of a tornado. Additionally, Ruginski et al. (2016) also reported on a distance heuristic when collecting damage estimates and think aloud statements for different types of hurricane cones of uncertainty. Thus, the distance heuristic seems to be a tool used to simplify the process of making a decision with uncertainty across different natural hazards. Interestingly, the distance was more often used in the numerical group. One hypothesis to explain this would be that participants were struggling to interpret the numerical uncertainty expressions and fell back on the distance heuristic to inform their decision-making.

Interesting regarding the boundary effect reported on above is that it apparently did not occur subconsciously. On the contrary, participants actively chose the *no damage* level if the house was located right outside the debris flow extent. This again supports the notion that participants struggled to interpret the uncertainty information, which is continuous and does not stop at a certain boundary. The boundary effect is further discussed in Chapter 5.5.4.

The last aspect, which must be addressed in the context of the decision-making strategy, is the potential influence of a risk aversion attitude presented in in Chapter 2.5. One participant explicitly described a strategy similar to this: *"I was more on the safe side. I would rather say that there could be a damage instead of not predicting any"* (translation by the author). This statement is hard to compare to the theoretical explanation of the risk aversion, which traditionally frames the concept through a gambling scenario with clear wins and losses. However, if the damages of a debris flow end up much larger than previously expected, this can be seen as a loss. Thus, the strategy to rather overestimate the damage instead of underestimating it, can be seen as a type of risk aversion. This could have caused an overestimation of damages in the scenario of study. The tendency towards making risk averse decisions could be tied to their personal attitude on risk taking as found by Kübler et al. (2020). Similar observations were made by Jenkins et al. (2017) when investigating different ways to communicate risk. Their participants explained that they took a *better to be safe than sorry* position when making decisions on evacuations. A risk aversion when making decisions with uncertainty was also found in the decision-making strategies of the study conducted by Korporaal et al. (2020). Another statement showing a risk aversion was already collected during the expert interviews. E5 mentioned that they would rather display a region in the higher danger zone in edge cases of hazard maps instead of choosing the lower one (e.g., red instead of blue) to deal with the responsibility from the political point of view. Hope and Hunter (2007) actually see a risk aversion attitude as rational, since "people are generally conservative and suffer from a loss more strongly than they value the corresponding gain".

The overall conclusion on the decision-making strategies is that only slight differences were found between the two communication methods. The uncertainty information represented through the colour scheme and the uncertainty legend were not as important as expected. On the contrary, the contour lines were extremely important for almost all participants. This shows that participants put uncertainty information to the side and focused on physical properties displayed on the map (e.g., the contour lines). The uncertainty was often a secondary source of information. Additionally, a series of cognitive shortcuts, mainly attribute substitution and a distance heuristic, were applied to simplify the complex decision-making process with uncertainty.

5.2. RQ1: Uncertainty Visualization

5.2.1. Influence of the Visualization Method on Damage Estimate

A significant difference between the damage estimates of the single-hue and the multi-hue stimuli was found in this study. The participants judged damages slightly higher with multi-hue stimuli compared to single-hue ones. This is in accordance with the expectations and previous studies, who found differences when comparing various uncertainty visualization methods (e.g., Ash et al. (2014), Kübler et al. (2020) and Leitner and Buttenfield, (2000)). However, the effect of this significance was minor, which weakens the statement above. If the weak significance is neglected, results of Cheong et al. (2016) would be confirmed as they found a slight difference between colour value and colour hue in terms of decision accuracy, but the difference was not significant.

Kübler et al. (2020) reported that participants in their study likely misinterpreted the uncertainty information regarding the danger zone when it was visualized through colour value. Participants interpreted lighter values to represent a lower danger zone. A similar misinterpretation could have occurred in this study: Participants interpreted the lighter colours in the single-hue method to represent a lower level of damage, when in reality the lighter values represented a higher uncertainty in the debris flow modelling. This effect was also observed in the results of the multi-hue stimuli. When interpreting the relative distribution of damage estimates, it was found that low damage levels were more often chosen for single-hue and high damage levels were more frequently chosen for multi-hue stimuli. Based on the finding that participants matched the uncertainty to the damage level, this could mean that participants were more risk averse (as discussed above) when making decisions with multi-hue stimuli. Thus, they would rather judge damages too high instead of too low. Investigating the correlation between the damage estimate and the uncertainty value in the context of RQ3 also showed a very slight difference in the strength of the relationships. For the single-hue stimuli the Spearman's correlation coefficient r_s resulted in a value of 0.868, while it was 0.886 for multi-hue stimuli (see Table 18). Thus, it could be argued that the slightly weaker relationship for the single-hue stimuli would mean that participants understood uncertainty more correctly with this method. However, the differences are certainly too small to be proof of a meaningful trend. It can be concluded from these observations that, no matter how the uncertainty is visualized, the non-expert audience had trouble understanding the concept of uncertainty.

5.2.2. Influence of the Visualization Method on Response Time

Since the response time did not significantly differ between the two visualization methods, it can be concluded that both methods required a similar amount of cognitive effort. This goes along with the expectations and could be because both methods made use of one or more properties of colour. Additionally, this reproduced results found by Cheong et al. (2016). One bias in this regard could be caused by the fact that participants did not know that they would be presented with more than one visualization method. Thus, it is possible that some participants did not realize that the colour scheme changed halfway throughout the trial, which was mentioned by two participants in the post-test. Irrespective of whether participants noticed the change or not, it can be concluded that the two visualization methods did not have an effect on the response time. However, it must also be mentioned that the range of the response times recorded in this study were rather large. This could be caused by the varying experiences and the different decision-making strategies among the participants. Yet, the ranges were similar across both visualization methods.

5.2.3. Influence of the Visualization Method on Eye Movements

Only slight differences in the eye movement patterns were found between the two visualization methods. Contrary to the expectations, participants did not spend a lot of time looking at the colour scheme overall. This could be due to the occurrence of a learning effect throughout the experiment. Thus, the participants had already memorized the colour scheme in later trials. Surprisingly, the decision-making strategies showed that the colour schemes were very important for participants. Thus, it could be hypothesized that participants were processing the colour scheme directly on the map display, where a lot of fixation time was registered, instead of looking at the colour scheme in the map legend. This might have been facilitated by the sequential nature of both colour schemes. Nevertheless, significantly more fixation time and visits were registered in the uncertainty legend for the multi-hue colour scheme, which could be due to the presence of multiple colours. But the effect size of this significance was rather low.

5.2.4. Post-Test: Preference Tasks

The two preference tasks posed to participants showed that the multi-hue colour scheme was strongly preferred both in terms of its suitability for decision-making as well as its aesthetics. This goes in line with the findings of Cheong et al. (2016), Klockow-McClain et al. (2020) or Miran et al. (2019), where the spectral colour hue method outperformed all others in the preference task. The main reason to prefer the multi-hue colour scheme was related to the contrast of the different hues, which enabled participants to distinguish between different colour levels. With the single-hue colour scheme they criticized that different brightness levels were hard to distinguish and the contrast with the base map was too low. These findings again reflect previous results by Miran et al. (2019). They also found that the higher contrast of a multi-hue colour scheme was the reason for its high rating in the preference task. It can be concluded that, in line with the expectations, participants perceived the multi-hue stimuli more suitable for the decision-making task.

From a theoretical perspective, the main argument to use the single-hue method is related to the uncertainty evoking properties. This is likely the reason why participants preferring the single-hue method thought it was intuitive. Yet, the majority of the participants appreciated that the uncertainty levels in the multi-hue colour scheme were more easily distinguishable. This sheds a light on a potential conflict between what is proposed by uncertainty visualization theory and what makes uncertainty visualizations readable to non-experts.

In terms of the colour chosen for the two colour schemes, a total of five participants mentioned that they would have chosen a different colour than green. They justified this opinion by saying that green did not convey danger such as the colour of red for instance. However, as humans have a hard time distinguishing different values of red (Katsnelson, 2021), a red-based colour scheme might not have performed better.

Even though there was a strong preference towards the multi-hue colour scheme, damage estimates and response times did not differ between the two visualization methods. Clearly, the preference for a certain method does not directly relate to the effectiveness of it (Cheong et al., 2016). However, when it is one's goal to create a map display, which is intuitive and inviting to read, the results of this preference task could be taken into account for the production of similar map displays in the future.

To sum up, the visualization method weakly influenced the decision outcome and did not affect the response time. Therefore, one could conclude that it does not matter which of the two colour schemes are used. However, it can be argued that the results found in the preference tasks should be considered if these colour schemes are further investigated. Visualizations incorporating uncertainty are only just emerging, especially in visualizations targeted at the public. It is certainly beneficial to use a popular colour scheme to gain the interest and acceptance of non-experts. In conclusion, the sequential multi-hue colour scheme presented in this study is seen as a suitable method to further investigate.

5.3. RQ2: Uncertainty Communication

5.3.1. Influence of the Communication Method on Damage Estimate

Against the expectations, the way uncertainty was communicated through the map legend did not have an influence on the damage estimates. This could be traced back to the lack of consideration of the map legend during the decision-making as uncovered by the eye tracking results. Again, slight differences in the relative distribution of damage estimates were found. The numerical group was more likely to choose higher damage estimates such as 5 – *substantial damage* and 7 –

very severe damage compared to the verbal group. This could mean that the numerical group was more risk averse. The overall lack of a significant difference reinforced findings summarized in the review by Wallsten and Budescu (1995), who found that the quality of decision outcomes was not affected by the uncertainty communication method in text-based studies.

Additionally, the tendency to map the uncertainty values to the damage estimate was almost equal between the two groups. The difference in the vagueness of the uncertainty expressions did not lead to a difference in how potential damages of debris flows were judged. Unfortunately, inconsistencies in how the uncertainty was interpreted with the two communication formats as investigated by Budescu et al. (2012, 2009) were not detectable due to the study design.

5.3.2. Influence of the Communication Method on Response Time

Just as expected, the verbal groups took significantly longer to make decisions compared to the numerical group. However, the rather low effect size of this difference must be considered. A potential cause of this difference could be the way the vague verbal expressions were interpreted. This could confirm the finding by Doyle et al. (2014) that participants tried to match the verbal expression to a precise numerical estimate before making their decision. This would increase cognitive load expressed through the higher response time. A more trivial hypothesis for why the verbal group took longer to answer would be that the longer uncertainty labels required more time to be read compared to the numerical values. As no other map-based study investigating these two uncertainty communication formats was found, comparisons to previous results were not possible. Additionally, reporting on response time is often missing in text-based uncertainty communication studies. Only Wallsten and Budescu (1995) reported on a previous study, where the response time did not differ between making a decision with numerical or verbal expressions.

5.3.3. Influence of the Communication Method on Eye Movements

The eye tracking results showed that the communication method significantly influenced the metrics recorded in various AOIs. The differences observed in the uncertainty label and uncertainty legend AOIs can be considered most relevant when taking the effect sizes into account (see Table 15). The differences in the total duration of fixations could be traced back to the length of the uncertainty labels for the verbal group. However, the longer labels would not lead to a higher number of visits, since this measure describes how often an AOI is entered (Tobii Pro AB, 2014). The significant difference in the number of visits in these AOIs shows that participants in the verbal group moved their gaze to and from the uncertainty label AOI more frequently compared to the numerical group. Thus, the cognitive effort to make the decision must have been higher for the verbal group. The pattern observed with the number of visits on the Likert scale AOIs was in accordance with the relative distribution of damage estimates recorded by the two communication groups. This shows that eye tracking proved to be a suitable source of data to investigate decision outcome as well as the decision-making process.

5.3.4. Post-Test: Importance of Map Legend, Trust & Task Difficulty

Participants rated the importance of the map legend high independently of their communication group. This stands in conflict with the results obtained from the overall eye movement patterns, which showed that the map legend was mostly neglected. This conflict again shows that what humans think they do – include the map legend for their decision-making – is not what they actually do – focus their gaze on the map display.

Another aspect related to RQ2 in the post-test was a difference in the trust put into the displayed information. Participants in the numerical group rated their trust higher compared to the verbal group, but the difference was not significant. This leads to the conclusion that the numerical uncertainty communication format conveys a slightly higher level of trust compared to vague verbal ones. This finding could be tied to the overall opinion that humans prefer to receive uncertainty information in the numerical form (Erev and Cohen, 1990). Jenkins et al. (2017) have argued that numerical expressions could be perceived as too precise and thus lead to a lower level of trust. This hypothesis was not confirmed by the observations in this study. Yet, the overall high rating of trust could support the common assumption that participants are confident in their decision-making when provided with uncertainty information (Kinkeldey et al., 2017).

Although both communication groups rated the task to be rather easy, a slight – yet statistically insignificant – difference was found. Difficulty ratings were lower for the numerical group compared to the verbal one. This is against the expectations posed above. Consequently, it can be argued that even though the vagueness of verbal expressions is seen to convey the nature of uncertainty more closely (Joslyn and Savelli, 2021), the preference to receive uncertainty in a numerical form (Erev and Cohen, 1990) could have guided the ratings of task difficulty.

It can be concluded that the two uncertainty communication methods did not influence the damage estimates in this study. This means that the uncertainty expressions were similarly interpreted. However, the effects of the communication methods could also have been overshadowed by the challenge of understanding the uncertainty information correctly. The combination of the response time and the eye movement patterns showed that that the processing of verbal uncertainty information required slightly more cognitive effort. Additionally, the perceived precision of numerical uncertainty expressions led to slightly higher levels of trust compared to the vague verbal ones. Thus, the numerical communication method can be seen as slightly superior. The same was argued in the recent review of text-based uncertainty communication studies by Dhami and Mandel (2022).

5.4. RQ1 & RQ2: Interaction between Visualization & Communication Method

The analysis showed that no significant interaction occurred regarding the damage estimate. When it comes to the response time, the picture is a bit more complicated. While the transition from the multi-hue colour scheme to the single-hue one led to an increase in response time for the numerical group, the response time was reduced for the verbal groups. This is the reason why, even though the LMM resulted in a significant interaction, none of the contrasts investigated in the post-hoc analysis were significant. Based on this analysis the aggregation of the data to answer RQ1 and RQ2 were considered non-problematic, as the variables did not significantly or consistently interact.

5.5. RQ3: Influence of Additional Information

5.5.1. Influence of Single Parameters on Damage Estimate

Besides the uncertainty value, the slope at the house location as well as the distance between the house and the riverbed were implicitly presented on the map stimuli. All three of these parameters could have guided the decision-making of participants.

The weakest relationship between one of the parameters and the damage estimates was found with the slope. The intuitive expectation was that a debris flow causes more severe damages in steep areas compared to gentle ones. The median damage estimates for the three slope categories follow this trend. However, participants chose a variety of different damage levels at all slope categories, which is expressed through the large difference between the 25th and the 75th percentile in Figure 52. A stronger relationship was found between the distance and the damage estimate. Additionally, the LMM calculated a strong effect of the distance. Thus, the presence of a distance heuristic as proposed above was confirmed. The strongest relationship was found between the uncertainty value and the damage estimate. This resulted in the strong matching of the uncertainty values to the damage Likert scale and proves the misinterpretation of the uncertainty information as elaborated at the beginning of the discussion.

In summary, the LMMs showed that damage estimates were most strongly affected by the distance and the uncertainty value. The slope at the house location only had a small to medium effect on the decisions. This could be due to the fact that extracting the slope from the map display was likely more complicated and required more profound map reading skills compared to the distance and the uncertainty value.

5.5.2. Influence of Parameter Interactions on Damage Estimate

A significant interaction was found between the slope and the distance category regarding the damage estimates. Based on the interaction plot in Figure 54 it was found that the damage estimates across different slope levels were rather consistent. This is especially the case at short distances. The parameter which seemed to have led to the spreading of damage estimates across the Likert scale was the distance. The clearest influence of distance were found when combined with a gentle and a steep slope. These patterns clearly show that the distance heuristic mentioned above was a strong predictor of the damage estimates in this study.

Slope and uncertainty value categories were also found to interact significantly in terms of the damage estimates. The interaction plot in Figure 55 showed that the influence of the slope category was again rather low, as damage estimates were rather similar between the slope levels. A negative relationship existed between the slope and the damage estimate at high and low uncertainty locations. This is rather counter intuitive considering the decision-making strategy descriptions in Chapter 5.1.3. Participants explained that they judged damages to be higher if the slope increased. This was, however, only the case at medium uncertainty locations. These observations shed a light on the potential weakness of self-reports. The uncertainty heuristic strongly influenced the resulting damage estimates and led to consistent estimates within each uncertainty category and across slope categories. However, within each slope category and across the uncertainty category, the uncertainty information led to a linear pattern in damage estimates. It can thus be concluded that the uncertainty value dominated the interaction and led to the decrease in damage estimates with an increasing uncertainty. This of course in accordance with the effect sizes of the individual parameters reported in the subchapters of 4.5.2.

Due to the presence of collinearity, it was not possible to investigate the interaction of the distance and the uncertainty value. This analysis would have certainly been interesting to get insight into the importance of the two heuristics associated with these parameters.

5.5.3. Importance of Additional Information

Clearly, the inclusion of all three parameters into a single LMM would have generated interesting results and given a clearer picture of which of the three sources of information were most important to predict the damage estimates. Nevertheless, the results would have likely not been reliable with the data collected in this study due to the high correlation between the distance and the uncertainty just mentioned. This is an aspect which should be more carefully controlled for in advance of future research.

Overall, it was observed that the heuristics relating to the distance and the uncertainty value dominated the decision-making outcomes. The influence of the slope information was a far weaker determinator for the damage estimates. This is partially in line with the expectations, since the uncertainty information guided decision-making most strongly and the additional information was also valuable. However, the difference between the influence of uncertainty and distance compared to the slope was certainly larger than expected.

When asked how they would improve the map displays, 14 out of the 43 participants suggested to include additional context information such as vegetation, infrastructure or spatial extents of past events. The question of whether the depiction of even more information would have supported a more successful and informed decision-making remains unanswered.

The fact that other information provided through the map displays was important in the decision-making process reproduces results of previous studies. For instance, Kübler et al. (2020) found that, while the cartographic design was important in their house-buying scenario, the house location was the most determining factor in the decision-making strategies. The danger zone and its associated uncertainty were less important. Additionally, Korporaal et al. (2020) reported that the uncertainty regarding the terrain slope was less important in their helicopter-landing scenario compared to the distance when participants made a decision between equally suitable sites. Unlike the results this thesis, these previous findings show that uncertainty information can actually be outweighed by other information during the decision-making process.

5.5.4. Boundary Effect

As presented in Chapter 4.5.4 a clear boundary effect was observed in the damage estimates for house locations at a high distance from the riverbed. This replicated previous results by Ash et al. (2014), Klockow-McClain et al. (2020) and Ruginski et al. (2016). It can be concluded that the boundary of the uncertainty surface significantly influenced how participants interpreted the potential damages. This links back to a heuristic of containment mentioned by Padilla et al. (2018), where the boundaries "help partition an information space into zones of relative semantic homogeneity". This heuristic explains why participants distinguished between a zone of damage inside the debris flow shape and a safe zone outside of it.

Closely considering the uncertainty information, this might come as a surprise, since the boundary of the uncertainty surface was arbitrarily set at an uncertainty level of 90%. Especially for the numerical uncertainty group, it was clearly visible that the debris flow extent only reached 90% uncertainty. Therefore, residual uncertainty was present right outside the displayed debris flow extent. However, these results further underline the fact that participants did not put much weight and attention on the map legend. Interestingly, one participant in the numerical group mentioned regrets regarding their strategy to judge the damage at locations outside the boundary. They stated that "*in hindsight I would trust the debris flow simulations less and critically question whether* *a debris flow would reach a house in cases where the house location was in an area where the modelled uncertainty was around 100% and choose a very minor damage instead of no damage"* (translation by the author).

It was proposed in Chapter 1.3.1.3 that a difference in the boundary effect could occur between the two colour schemes. Yet, even though the perceived fuzziness of the single-hue methods was higher, the effect sizes of the boundary effect of the two visualization methods were similar. Thus, the boundary effect can be observed no matter whether the boundary is displayed sharply or fuzzily. This is not in accordance with results by Ruginski et al. (2016), who found that the boundary heuristic was more dominant in a fuzzy visualization.

5.5.5. Post-Test: Importance of Contours

The post-test showed that the importance of the contour lines was very high. This was also reflected in the descriptions of the decision-making strategies regarding the damage estimates. However, the effect of this parameter was overridden by the distance and the uncertainty value when it comes to the collected results. Participants mentioned that the slope helped them to imagine the movement of the debris flow and get an overview of the terrain.

In conclusion, it was found that decisions were strongly guided by the uncertainty heuristic (attribute substitution) discussed above. However, the distance heuristic seemed to be almost as strong. Thus, damage estimates were strongly influenced by the uncertainty value and the distance through heuristics, which both served as ways to simplify decision-making with uncertainty. Although the contour lines were seen as an important aspect of the map stimuli, the slope had little effect on the damage estimates. In the context of the distance heuristic, a strong boundary effect was observed. This shows that the uncertainty boundary displayed in uncertainty depictions is a strong predictor of damage estimates across different natural hazards.

5.6. Limitations of the Study

The following limitations must be taken into account when reflecting on the work at hand: One limitation of this study is the rather small sample size, which was limited by the available time resources. The conduction of the study was highly uncertain as it took place in the midst of the Sars Cov-2 pandemic. The recruiting of participants was challenged by to the restrictions and measures which were in place during the time of the data collection. The results of the study would likely be more reliable if a larger sample had been tested. Nevertheless, the sample size of 43 participants was seen as a success considering the time constraints and the circumstances at the time.

One aspect Padilla et al. (2015) criticized in their own study was that their decision-making task did not have real-world consequences. It is always questionable how closely a study scenario resembles a real-world setting. The scenario and the task used in this study might not be realistic as the underestimation of the damage at a certain house did not have consequences for the decision-makers. Still, the decision-making task was informed by the experts interviews and conceptualized in a way that non-experts would be able to imagine the scenario easily. The discrepancy between a real-world decision and the one made in this study is inevitable.

Another limitation relating to the representativeness of the study concerns the uncertainty data. Although the uncertainty information was modelled using the RAMMS::DEBRIS FLOW (SLF/WSL, 2017a) which is used in practice to create simulations and hazard maps, the chosen workflow was rather simple. This can be traced back to the temporal constraints. Additionally, the overall goal of the study was to investigate ways to visualize and communicate uncertainty and not to perform an in-depth sensitivity analysis of the model.

The study design influenced the results as, for instance, answers from the post-test questionnaire could be analysed for the between-group variable (communication method) but not for the within-group variable (visualization method). Thus, some information on potential differences caused by the visualization methods might remain hidden in the data.

Lastly, two aspects made the comparison of the study at hand with previous research challenging. Firstly, the damage estimate is not a common type of decision outcome in uncertainty visualization research, which traditionally report on decision accuracy (Hullman et al., 2019). Yet, the interest of the author was not to declare a winning method but to investigate the effect of different uncertainty visualization and communication methods on map-based decision-making. Nevertheless, more direct comparisons with previous studies would have been interesting. Secondly, no map-based study has compared numerical and verbal uncertainty expressions to the knowledge of this author so far. Thus, some results presented in this thesis represent first findings in the investigated field and cannot be set into the context of other studies.

6. Conclusion & Outlook

This empirical, map-based study investigated two different visualization (single-hue and multi-hue) and communication (numerical and verbal) methods to convey uncertainty information inherently tied to predicting future debris flows. It studied differences in damage estimate, response time, eye movements as well as decision-making strategy. Additionally, it discussed whether uncertainty information was guiding the decisions or if other information in the map display were superior.

Regarding the influence of the two visualization methods (RQ1), it was found that participants judged damages significantly higher with the multi-hue compared to the single-hue method. Yet the effect size of this variable was low. This difference could be caused by the differing contrast between the two colour schemes. Response time was not affected by the visualization method. A different picture was drawn when testing the two ways of communicating uncertainty with numerical and verbal expressions in the map legend (RQ2). Unlike the expectations, the communication method did not have an influence on the damage estimate. However, it did lead to a significantly higher response time for the verbal group. This difference can be traced back to the amount of cognitive effort associated with the vagueness of verbal expressions and is in accordance with previous results. Investigating which sources of information – explicitly or inherently represented in the map displays – guided the decision-making process (RQ3), showed that the uncertainty information and the distance between the house location and the riverbed were the best predictors of damage estimate. The strong influence of these two parameters can be explained by the application of attribute substitution – causing a matching of the uncertainty information with the damage scale – and a distance heuristic. The presence of a strong boundary effect, as found in numerous previous studies, was confirmed in the context of the distance parameter. The slope represented through contour lines was only of secondary importance in estimating damage.

The analysis of the decision-making strategies certainly added value to the quantitative measures collected during the study. It became clear that decision-making strategies varied heavily between participants. However, no clear influence of the uncertainty communication method was found. Most importantly, the uncertainty information was not the only aspect to influence decision-making. Despite its low influence on damage estimates, the slope was a lot more relevant in the decision-making strategies than expected. Additionally, discrepancies regarding the importance of the colour scheme were uncovered when comparing strategy descriptions with eye tracking data. Participants perceived the colour scheme as highly important but did not really take it into account based on their eye movements. In conclusion, it was found that the decision-making process including uncertainty is just as complicated as expected. This is why heuristics such as attribute substitution, a distance heuristic and a boundary effect were observed.

Likely one of the most important findings made in this study is that non-expert participants had trouble correctly interpreting the uncertainty information. Participants essentially scaled the uncertainty values to the damage Likert scale to facilitate their decisions. This reinforces doubts of experts that the visualization of uncertainty could lead to a misunderstanding on the map reader's side. This finding must be taken into account when revisiting the results of each research question above. However, it is not considered an argument to give up on uncertainty visualization as a whole. On the contrary, it is a motivation to invest resources into training the public on the concept of uncertainty and how to incorporate it into their decision-making process. Due to its inherent presence, uncertainty must become part of a public discourse. Topics such as climate change and natural hazards certainly represent suitable platforms to do so.

In their review of uncertainty visualization studies, Hullman et al. (2019) stated that only 0.5% of studies in this field made use of eye tracking technology to collect data. Through the application of eye tracking, this study was able to uncover findings which would not have been deductible from the decision outcomes and the response times alone. By tracking the eye movements it was observed that the map legend was mostly neglected by participants, which could explain the weakness or lack of differences in damage estimates between the different visualization and communication methods. Therefore, the method of eye tracking was confirmed as a valuable tool in empirical cartographic research, especially in the realm of spatial decision-making.

The results of the study imply that design choices made in uncertainty visualization might be of secondary importance as long as profound understanding of the concept of uncertainty is not present. Varying the colour scheme in the maps (RQ1) did not have a clear influence on the damage estimates of participants. Furthermore, the different types of map legends (RQ2) only had a small effect on the study results. Cheong et al. (2016) interpreted the lack of significant differences in their study as follows: "This result might be interpreted to indicate that the cartography 'doesn't matter' in this instance of decision-making under uncertainty". The specific findings made in this thesis indicate that the profound understanding of uncertainty information is a clear precondition to finding the best fit to visualize uncertainty for a specific field of application.

This thesis leaves some research gaps untouched. It is considered valuable to further investigate the depiction of uncertainty specifically in debris flow predictions. Participants mentioned that the display of additional information such as infrastructure, vegetation and spatial extents of past events would have helped them to make their decisions. The post-test questionnaire also showed that a certain number of participants did not interpret the green colour schemes to represent danger or hazard. Future research could investigate single- and multi-hue colour schemes based on different colours. Additionally, it would be interesting to test a larger variety of established uncertainty visualization methods on debris flow prediction maps. In terms of the uncertainty communication format of the map legends, future research could investigate how a combination of the numerical and the verbal uncertainty, as suggested by Budescu et al. (2012, 2009), would perform in a map-based context. The uncertainty depictions could be further developed to include interactive features such as functions to zoom, to pan and to display different sources of information. A pioneer interactive application in the context visualizing uncertainty in natural hazards has already been developed by Kunz et al. (2011a). However, empirical results are still missing in this realm. Although Joslyn and Savelli (2021) have found that non-experts are capable to make decisions with uncertainty visualizations, this study has shown that non-experts struggle to understand uncertainty information. It would be interesting to test the applied methods with more specific groups of participants such as natural hazard experts or practitioners in this field. This would help to compare the results reported above with ones collected with experts. Insights into the understanding of the displayed uncertainty by experts could help to further investigate the causes of the misinterpretation of uncertain information reported on above.

In conclusion, this study has shown that non-experts struggle to interpret uncertainty but are likely not aware of it. The lack of familiarity with uncertainty information is a potential reason for this. Non-experts could get more used to the concept of uncertainty if this information was more often communicated to the public through maps. The author of this thesis argues that being able to correctly process uncertain information is a precondition to investigate how humans take uncertainty into account in map-based decision-making. Thus, resources should be mobilized to create profound training on dealing with uncertainty in specific fields of application. This would enable a more reliable, empirical evaluation of uncertainty visualizations. Over the past few decades, it has become evident that there is no *one size fits all* solution on creating a successful uncertainty visualization (Pang, 2008). Therefore, the value of studies such as the one at hand, which apply and test uncertainty visualization and communication methods to specific fields of application, must be acknowledged. To the knowledge of the author, this study was the first to empirically test uncertainty visualizations for debris flow predictions. Additionally, it was the first map-based study to compare numerical and verbal uncertainty expressions. Thus, the results generated in this thesis are seen as a valuable contribution to the fields of uncertainty visualization in natural hazard management.

7. Literature

Aerts, J.C.J.H., Clarke, K.C., Keuper, A.D., 2003. Testing popular visualization techniques for representing model uncertainty. Cartogr. Geogr. Inf. Sci. 30, 249–261. https://doi.org/10.1559/152304003100011180

Albaum, G., 1997. The Likert scale revisited: An alternate version. Int. J. Mark. Res. Soc. 39, 331–348. https://doi.org/10.1177/147078539703900202

ARE, FOEN, 2005. Empfehlung: Raumplanung und Naturgefahren, Naturgefahren, Federal Office for Spatial Development (ARE), Federal Office for the Environment (FOEN). Bern.

Ash, K.D., Schumann, R.L., Bowser, G.C., 2014. Tornado warning trade-offs: Evaluating choices for visually communicating risk. Weather. Clim. Soc. 6, 104–118. https://doi.org/10.1175/WCAS-D-13-00021.1

Balleine, B.W., 2007. The neural basis of choice and decision making. J. Neurosci. 27, 8159–8160. https://doi.org/10.1523/JNEUROSCI.1939-07.2007

Bertin, J., 1983. Semiology of graphics: diagrams, networks, maps. University of Wisconsin Press, Madison, Wis.

Bisantz, A.M., Cao, D., Jenkins, M., Pennathur, P.R., Farry, M., Roth, E., Potter, S.S., Pfautz, J., 2011. Comparing uncertainty visualizations for a dynamic decision-making task. J. Cogn. Eng. Decis. Mak. 5, 277–293. https://doi.org/10.1177/1555343411415793

Bisantz, A.M., Stone, R.T., Pfautz, J., Fouse, A., Farry, M., Roth, E., Nagy, A.L., Thomas, G., 2009. Visual representations of meta-information. J. Cogn. Eng. Decis. Mak. 3, 67–91. https://doi.org/10.1518/155534309X433726

Boone, H.N.J., Boone, D.A., 2012. Analyzing Likert data. J. Ext. 50, 1–5.

Borland, D., Taylor, R.M., 2007. Rainbow color map (still) considered harmful. IEEE Comput. Graph. Appl. 27, 14–17. https://doi.org/10.1109/MCG.2007.323435

Boukhelifa, N., Duke, D.J., 2009. Uncertainty visualization - Why might it fail?, in: Proceedings of the 27th International Conference on Human Factors in Computing Systems, CHI 2009. Boston, MA, USA, pp. 4051–4056. https://doi.org/10.1145/1520340.1520616

Brewer, C.A., 1994. Color use guidelines for mapping and visualization, in: MacEachren, A.M., Taylor, D.R.F. (Eds.), Modern Cartography Series. Elsevier Science Ltd, pp. 123–147. https://doi.org/10.1016/B978-0-08-042415-6.50014-4

Brewer, C.A., Harrower, M., The Pennsylvania State University, 2013. ColorBrewer 2.0: Color advice for cartography [WWW Document]. URL https://colorbrewer2.org/#

Brodlie, K., Osorio, R.A., Lopes, A., 2012. A review of uncertainty in data visualization, in: Dill, J., Earnshaw, R., Kasik, D., Vince, J., Wong, P. (Eds.), Expanding the Frontiers of Visual Analytics and Visualization. Springer, London, pp. 81–109. https://doi.org/10.1007/978-1-4471-2804-5_6

Brus, J., Popelka, S., Brychtova, A., Svobodova, J., 2012. Exploring effectiveness of uncertainty visualization methods by eye-tracking, in: Proceedings of the 10th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences. Florianopolis, Brazil, pp. 215–220.

Budescu, D. V., Broomell, S., Por, H.-H., 2009. Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. Psychol. Sci. 20, 299–308. https://doi.org/10.1111/j.1467-9280.2009.02284.x

Budescu, D. V., Por, H.-H., Broomell, S.B., 2012. Effective communication of uncertainty in the IPCC reports. Clim. Change 113, 181–200. https://doi.org/10.1007/s10584-011-0330-3

Buttenfield, B.P., 1993. Representing data quality. Cartographica 30, 1–7. https://doi.org/10.3138/232H-6766-3723-5114

Buttenfield, B.P., Ganter, J.H., 1990. Visualization and GIS: What should we see? What might we miss?, in: Proceedings of the 4th International Symposium on Spatial Data Handling. Zurich, Switzerland, pp. 307–316.

BWW, BRP, BUWAL, 1997. Naturgefahren: Empfehlungen 1997 Berücksichtigung der Hochwassergefahren bei raumwirksamen Tätigkeiten. Biel.

Cedilnik, A., Rheingans, P., 2000. Procedural annotation of uncertain information, in: Proceedings Visualization 2000. VIS 2000 (Cat. No.00CH37145). IEEE, Salt Lake City, UT, USA, pp. 77–84. https://doi.org/10.1109/visual.2000.885679

Cheong, L., Bleisch, S., Kealy, A., Tolhurst, K., Wilkening, T., Duckham, M., 2016. Evaluating the impact of visualization of wildfire hazard upon decision-making under uncertainty. Int. J. Geogr. Inf. Sci. 30, 1377–1404. https://doi.org/10.1080/13658816.2015.1131829

Chesneau, E., Ruas, A., Bonin, O., 2005. Colour contrasts analysis for a better legibility of graphic signs on risk maps. 22nd Int. Cartogr. Conf. 1–10.

Clapham, S.B., 1992. A formal approach to the visualization of spatial data quality, in: GIS/LIS '92 Annual Conference and Exposition. San Jose, California, pp. 138–149.

Clark, L.A., Watson, D., 2019. Constructing validity: New developments in creating objective measuring instruments. Psychol. Assess. 31, 1412–1427. https://doi.org/10.1037/pas0000626

Çöltekin, A., Heil, B., Garlandini, S., Fabrikant, S.I., 2009. Evaluating the effectiveness of interactive map interface designs: A case study integrating usability metrics with eye-movement analysis. Cartogr. Geogr. Inf. Sci. 36, 5–17. https://doi.org/10.1559/152304009787340197

Correll, M., Moritz, D., Heer, J., 2018. Value-suppressing uncertainty palettes, in: Conference on Human Factors in Computing Systems - Proceedings. Montral, QC, Canada, pp. 1–11. https://doi.org/10.1145/3173574.3174216

Cox, J., House, D., Lindell, M., 2013. Visualizing uncertainty in predicted hurricane tracks. Int. J. Uncertain. Quantif. 3, 143–156. https://doi.org/10.1615/int.j.uncertaintyquantification.2012003966

Deitrick, S., 2012. Evaluating implicit visualization of uncertainty for public policy decision support, in: Proceedings AutoCarto 2012. Columbus, Ohio, USA.

Deitrick, S., Edsall, R., 2006. The influence of uncertainty visualization on decision making: An empirical evaluation, in: Riedl, A., Kainz, W., Emes, G.A. (Eds.), Progress in Spatial Data Handling. Springer, Berlin, Heidelberg, pp. 719–738. https://doi.org/10.1007/3-540-35589-8_45

Deubelbeiss, Y., Graf, C., 2013. Two different starting conditions in numerical debris flow models – Case study at Dorfbach, Randa (Valais, Switzerland), in: Graf, C. (Ed.), Mattertal - Ein Tal in Bewegung. Eidg. Forschungsanstalt für Wald, Schnee und Landschaft WSL, pp. 125–138.

Dhami, M.K., Mandel, D.R., 2022. Communicating uncertainty using words and numbers. Trends Cogn. Sci. 26, 514–526. https://doi.org/10.1016/j.tics.2022.03.002

Doyle, E.E.H., McClure, J., Paton, D., Johnston, D.M., 2014. Uncertainty and decision making: Volcanic crisis scenarios. Int. J. Disaster Risk Reduct. 10, 75–101. https://doi.org/10.1016/j.ijdrr.2014.07.006

Drecki, I., 2002. Visualisation of uncertainty in geographical data, in: Shi, W., Fisher, P.F., Goodchild, M.F. (Eds.), Spatial Data Quality. Taylor & Francis, London, pp. 140–159. https://doi.org/10.1201/b12657

Ehrlinger, J., Readinger, W.O., Kim, B., 2016. Decision-making and cognitive biases. Encycl. Ment. Heal. 12, 5–12. https://doi.org/10.1016/B978-0-12-397045-9.00206-8

Epstein, S., 1994. Integration of the cognitive and the psychodynamic unconscious. Am. Psychol. 49, 709–724. https://doi.org/10.1037/0003-066x.49.8.709

Erev, I., Cohen, B.L., 1990. Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. Organ. Behav. Hum. Decis. Process. 45, 1–18. https://doi.org/10.1016/0749-5978(90)90002-Q

Fabrikant, S.I., Hespanha, S.R., Hegarty, M., 2010. Cognitively inspired and perceptually salient graphic displays for efficient spatial inference making. Ann. Assoc. Am. Geogr. 100, 1–17. https://doi.org/10.1080/00045600903362378

Field, A., Miles, J., Field, Z., 2012. Discovering statistics using R. SAGE Publications.

Fischhoff, B., 2012. Communicating uncertainty: Fulfilling the duty to inform. Issues Sci. Technol. 28, 63–70.

Fisher, P.F., 1994. Randomization and sound for the visualization of uncertain spatial information, in: Unwin, I.D., Hearnshaw, H. (Eds.), Visualization in Geographic Information Systems. John Wiley & Sons, London, UK, pp. 181–185.

Fisher, P.F., 1993. Visualizing uncertainty in soil maps by animation. Cartographica 30, 20–27. https://doi.org/10.3138/B204-32P4-263L-76W0

FOEN, 2016. Schutz vor Massenbewegungsgefahren: Vollzugshilfe für das Gefahrenmanagement von Rutschungen, Steinschlag und Hangmuren, Umwelt-Vollzug, Federal Office for the Environment (FOEN). Bern.

FOEN, 2015. Was sagen Gefahrenkarten aus?, Federal Office for the Environment (FOEN).

Frank, F., Huggel, C., McArdell, B.W., Vieli, A., 2019. Landslides and increased debris-flow activity: A systematic comparison of six catchments in Switzerland. Earth Surf. Process. Landforms 44, 699–712. https://doi.org/10.1002/esp.4524

Galley, N., Betz, D., Biniossek, C., 2015. Fixation durations - Why are they so highly variable? pp. 1–26.

Geoportal des Kantons Bern, 2019. Naturgefahrenkarten 1:5'000 des Kantons Bern [WWW Document]. URL https://www.map.apps.be.ch/pub/synserver?project=a42pub_gk5&x=2638895.375031751&y=1165169.908729152&scale=5000&rotation=0&view=Ansicht_Standard&basemapview=HK_AV_grau&client=core&language=de (accessed 4.6.22).

Gershon, N., 1998. Visualization of an imperfect world. IEEE Comput. Graph. Appl. 18, 43–45. https://doi.org/10.1109/38.689662

Gong, P., Chen, J., 1992. Boundary uncertainties in digitized maps I: Some possible determination methods, in: GIS/LIS '92 Annual Conference and Exposition. San Jose, California, pp. 274–281.

Graf, C., Christen, M., McArdell, B.W., Bartelt, P., 2019. An overview of a decade of applied debris-flow runout modeling in Switzerland: Challenges and recommendations, in: Kean, J.W., Coe, J.A., Santi, P.M., Guillen, B.K. (Eds.), Association of Environmental and Engineering Geologists Special Publication: Vol. 28. Debris-Flow Hazards Mitigation: Mechanics, Monitoring, Modeling, and Assessment. Association of Environmental and Engineering Geologists, pp. 685–692.

Harold, J., Lorenzoni, I., Shipley, T.F., Coventry, K.R., 2016. Cognitive and psychological science insights to improve climate change data visualization. Nat. Clim. Chang. 6, 1080–1089. https://doi.org/10.1038/nclimate3162

Harrower, M., 2003. Representing uncertainty: Does it help people make better decisions?

Hegarty, M., Smallman, H.S., Stull, A.T., 2012. Choosing and using geospatial displays: Effects of design on performance and metacognition. J. Exp. Psychol. Appl. 18, 1–17. https://doi.org/10.1037/a0026625

Hirschberg, J., Badoux, A., McArdell, B.W., Leonarduzzi, E., Molnar, P., 2021a. Limitations of rainfall thresholds for debris-flow prediction in an Alpine catchment. Nat. Hazards Earth Syst. Sci. Discuss. 1–26. https://doi.org/10.5194/nhess-2021-135

Hirschberg, J., Fatichi, S., Bennett, G.L., McArdell, B.W., Peleg, N., Lane, S.N., Schlunegger, F., Molnar, P., 2021b. Climate change impacts on sediment yield and debris-flow activity in an Alpine catchment. J. Geophys. Res. Earth Surf. 126, 1–26. https://doi.org/10.1029/2020JF005739

Hitz, O.M., Amacher, M., Künzi, R., Schwab, S., 2014. Hochwasserschutz und wasserbauliche Planung im Wandel der Zeit : Beispiel Gryfenbach (Lauterbrunnen), in: Tagungsband Internationales Symposium: Wasser- Und Flussbau Im Alpenraum. Versuchsanstalt für Wasserbau, Hydrologie und Gaziologie, ETH Zürich, Zürich, pp. 603–613.

Hoffrage, U., Gigerenzer, G., 1998. Using natural frequencies to improve diagnostic inferences. Acad. Med. 73, 538–540. https://doi.org/10.1097/00001888-199805000-00024

Hope, S., Hunter, G.J., 2007. Testing the effects of thematic uncertainty on spatial decision-making. Cartogr. Geogr. Inf. Sci. 34, 199–214. https://doi.org/10.1559/152304007781697884

Howard, D., MacEachren, A.M., 1996. Interface design for geographic visualization: Tools for representing reliability. Cartogr. Geogr. Inf. Syst. 23, 59–77. https://doi.org/10.1559/152304096782562109

Huggel, C., Kääb, A., Haeberli, W., Krummenacher, B., 2003. Regional-scale GIS models for assessment of hazards from glacier lake outbursts: Evaluation and application in the Swiss Alps. Nat. Hazards Earth Syst. Sci. 3, 647–662. https://doi.org/10.5194/nhess-3-647-2003

Hullman, J., 2020. Why authors don't visualize uncertainty. IEEE Trans. Vis. Comput. Graph. 26, 130–139. https://doi.org/10.1109/TVCG.2019.2934287

Hullman, J., Qiao, X., Correll, M., Kale, A., Kay, M., 2019. In pursuit of error: A survey of uncertainty visualization evaluation. IEEE Trans. Vis. Comput. Graph. 25, 903–913. https://doi.org/10.1109/TVCG.2018.2864889

Hunter, G.J., 2005. Managing uncertainty in GIS, in: Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W. (Eds.), Geographical Information Systems: Principles, Techniques, Management and Applications. pp. 633–641.

IPCC, 2014. Klimaänderung 2014: Synthesebericht. Beitrag der Arbeitsgruppen I, II und III zum Fünften Sachstandsbericht des Zwischenstaatlichen Ausschusses für Klimaänderungen (IPCC). Genf, Schweiz.

IPCC, 2013. Summary for Policymakers, in: Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., Midgley, P.M. (Eds.), Chlimate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

IPCC WGI Technical Support Unit, 2018. IPCC Visual Style Guide for Authors.

Irwin, D.E., 2004. Fixation location and fixation duration as indices of cognitive processing, in: Henderson, J.M., Ferreira, F. (Eds.), The Integration of Language, Vision, and Action: Eye Movements and the Visual World. Psychology Press, New York, pp. 105–135.

Itti, L., Koch, C., Niebur, E., 1998. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. Pattern Anal. Mach. Intell. 20, 1254–1259. https://doi.org/10.1109/34.730558

Jakob, M., Stein, D., Ulmi, M., 2012. Vulnerability of buildings to debris flow impact. Nat. Hazards 60, 241–261. https://doi.org/10.1007/s11069-011-0007-2

Jebb, A.T., Ng, V., Tay, L., 2021. A review of key Likert scale development advances: 1995–2019. Front. Psychol. 12, 1–14. https://doi.org/10.3389/fpsyg.2021.637547

Jenkins, S.C., Harris, A.J.L., Lark, R.M., 2017. Maintaining credibility when communicating uncertainty: The role of communication format, in: Proceedings of the 39th Annual Meeting of the Cognitive Science Society. Cogntivie Science Society, pp. 582–587.

Johannsen, I.M., Fabrikant, S.I., Evers, M., 2018. How do texture and color communicate uncertainty in climate change map displays?, in: 10th International Conference on Geographic Information Science. Leibniz, pp. 1–6. https://doi.org/10.4230/LIPIcs.GIScience.2018.37

Joshi, A., Kale, S., Chandel, S., Pal, D.K., 2015. Likert scale: Explored and explained. Br. J. Appl. Sci. Technol. 7, 396–403. https://doi.org/10.9734/bjast/2015/14975

Joslyn, S., LeClerc, J., 2013. Decisions with uncertainty: The glass half full. Curr. Dir. Psychol. Sci. 22, 308–315. https://doi.org/10.1177/0963721413481473

Joslyn, S., Pak, K., Jones, D., Pyles, J., Hunt, E., 2007. The effect of probabilistic information on threshold forecasts. Weather Forecast. 22, 804–812. https://doi.org/10.1175/WAF1020.1

Joslyn, S., Savelli, S., 2021. Visualizing uncertainty for non-expert end users: The challenge of the deterministic construal error. Front. Comput. Sci. 2, 1–12. https://doi.org/10.3389/fcomp.2020.590232

Joslyn, S.L., LeClerc, J.E., 2012. Uncertainty forecasts improve weather-related decisions and attenuate the effects of forecast error. J. Exp. Psychol. Appl. 18, 126–140. https://doi.org/10.1037/a0025185

Just, M.A., Carpenter, P.A., 1976. Eye fixations and cognitive processes. Cogn. Psychol. 8, 441–480. https://doi.org/10.1016/0010-0285(76)90015-3

Kahnemann, D., Slovic, P., Tversky, A. (Eds.), 1982. Judgement under uncertainty: Heuristics and biases. Cambridge University Press. https://doi.org/10.1017/CBO9780511809477

Kahnemann, D., Tversky, A., 1982. Subjective probability: A judgement of representativeness, in: Kahnemann, D., Slovic, P., Tversky, A. (Eds.), Judgement under Uncertainty: Heuristics and Biases. Cambridge University Press, pp. 32–47. https://doi.org/10.1017/CBO9780511809477

Kandlikar, M., Risbey, J., Dessai, S., 2005. Representing and communicating deep uncertainty in climate-change assessments. Comptes Rendus Geosci. 337, 443–455. https://doi.org/10.1016/j.crte.2004.10.010

Katsnelson, A., 2021. Fixing figures for colour blindness: Images can be made more accessible by choosing hues, shapes and textures carefully. Nature 598, 224–225.

Kinkeldey, C., MacEachren, A.M., Riveiro, M., Schiewe, J., 2017. Evaluating the effect of visually represented geodata uncertainty on decision-making: Systematic review, lessons learned, and recommendations. Cartogr. Geogr. Inf. Sci. 44, 1–21. https://doi.org/10.1080/15230406.2015.1089792

Kinkeldey, C., MacEachren, A.M., Schiewe, J., 2014. How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. Cartogr. J. 51, 372–386. https://doi.org/10.1179/1743277414Y.0000000099

Klockow-McClain, K.E., McPherson, R.A., Thomas, R.P., 2020. Cartographic design for improved decision making: Trade-offs in uncertainty visualization for tornado threats. Ann. Am. Assoc. Geogr. 110, 314–333. https://doi.org/10.1080/24694452.2019.1602467

Koehler, K.L., Guo, F., Zhang, S., Eckstein, M.P., 2014. What do saliency models predict? J. Vis. 14, 1–27. https://doi.org/10.1167/14.3.14

Korporaal, M., Ruginski, I.T., Fabrikant, S.I., 2020. Effects of uncertainty visualization on map-based decision making under time pressure. Front. Comput. Sci. 2, 1–20. https://doi.org/10.3389/fcomp.2020.00032

Krassanakis, V., 2013. Exploring the map reading process with eye movement analysis, in: Proceedings of the 1st International Workshop on Eye Tracking for Spatial Research (in Conjunction with COSIT 2013). Scarborough, United Kingdom, pp. 2–7.

Kubicek, P., Sasinka, C., 2011. Thematic uncertainty visualization usability - comparison of basic methods. Ann. GIS 17, 253–263. https://doi.org/10.1080/19475683.2011.625978

Kubicek, P., Sasinka, C., Stachon, Z., 2012. Uncertainty visualization testing, in: 4th International Conference on Cartography and GIS, Vol. 1. Albena, Bulgaria, pp. 247–256.

Kübler, I., Richter, K.F., Fabrikant, S.I., 2020. Against all odds: Multicriteria decision making with hazard prediction maps depicting uncertainty. Ann. Am. Assoc. Geogr. 110, 661–683. https://doi.org/10.1080/24694452.2019.1644992

Kunz, M., Grêt-Regamey, A., Hurni, L., 2011a. Customized visualization of natural hazards assessment results and associated uncertainties through interactive functionality. Cartogr. Geogr. Inf. Sci. 38, 232–242. https://doi.org/10.1559/15230406382232

Kunz, M., Grêt-Regamey, A., Hurni, L., 2011b. Visualization of uncertainty in natural hazards assessments using an interactive cartographic information system. Nat. Hazards 59, 1735–1751. https://doi.org/10.1007/s11069-011-9864-y

Kusumastuti, S.A., Pollard, K.A., Oiknine, A.H., Dalangin, B., Raber, T.R., Files, B.T., 2022. Practice improves performance of a 2D uncertainty integration task within and across visualizations. IEEE Trans. Vis. Comput. Graph. 1–12. https://doi.org/10.1109/TVCG.2022.3173889

Leitner, M., Buttenfield, B.P., 2000. Guidelines for the display of attribute certainty. Cartogr. Geogr. Inf. Sci. 27, 3–14. https://doi.org/10.1559/152304000783548037

Li, X., Cöltekin, A., Kraak, M.-J., 2010. Visual exploration of eye movement data using the space-time-cube, in: Fabrikant, S.I., Reichenbach, T., van Kreveld, M., Schlieder, C. (Eds.), Geographic Information Science. GIScience 2010. Lecture Notes in Computer Science, Vol 6292. Springer, Berlin, Heidelberg, pp. 295–309. https://doi.org/10.1007/978-3-642-15300-6_21

Liechti, K., Matter, D., Lustenberger, F., Badoux, A., 2022. Unwetterschäden in der Schweiz 2021: Hochwasser, Murgänge, Rutschungen und Sturzprozesse. Wasser, Energie, Luft 114, 85–93.

Likert, R., 1932. A technique for measurement of attitudes. Arch. Psychol. 22, 5–55.

LimeSurvey GmbH, 2021. LimeSurvey - An Open Source survey tool: Community Edition Version 3.27.23.

Liu, L., Boone, A.P., Ruginski, I.T., Padilla, L., Hegarty, M., Creem-Regehr, S.H., Thompson, W.B., Yuksel, C., House, D.H., 2017. Uncertainty visualization by representative sampling from prediction ensembles. IEEE Trans. Vis. Comput. Graph. 23, 2165–2178. https://doi.org/10.1109/TVCG.2016.2607204

Loucks, D.P., van Beek, E., Stedinger, J.R., Dijkman, J.P.M., Villars, M.T., 2005. Model sensitivity and uncertainty analysis, in: Water Resources Systems Planning and Management: An Introduction to Methods, Models and Applications. UNESCO, pp. 254–290.

MacEachren, A.M., 1995. How maps work: Representation, visualization, and design. The Guilford Press, New York.

MacEachren, A.M., 1992. Visualizing uncertain information. Cartogr. Perspect. 13, 10–19. https://doi.org/10.14714/CP13.1000

MacEachren, A.M., Brewer, C.A., Pickle, L.W., 1998. Visualizing georeferenced data: Representing reliability of health statistics. Environ. Plan. A 30, 1547–1561. https://doi.org/10.1068/a301547

MacEachren, A.M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., Hetzler, E., 2005. Visualizing geospatial information uncertainty: What we know and what we need to know. Cartogr. Geogr. Inf. Sci. 32, 139–160. https://doi.org/10.1559/1523040054738936

MacEachren, A.M., Roth, R.E., O'Brien, J., Li, B., Swingley, D., Gahegan, M., 2012. Visual semiotics & uncertainty visualization: An empirical study. IEEE Trans. Vis. Comput. Graph. 18, 2496–2505. https://doi.org/10.1109/TVCG.2012.279

Magezi, D.A., 2015. Linear mixed-effects models for within-participant psychology experiments: An introductory tutorial and free, graphical user interface (LMMgui). Front. Psychol. 6, 1–7. https://doi.org/10.3389/fpsyg.2015.00002

Martin, D.W., 2008. Doing psychology experiments, 7th ed. Thomson Higher Education, Bemont, CA.

Mastrandrea, M.D., Field, C.B., Stocker, T.F., Edenhofer, O., Ebi, K.L., Frame, D.J., Held, H., Kriegler, E., Mach, K.J., Matschoss, P.R., Plattner, G.-K., Yohe, G.W., Zwiers, F.W., 2010. Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties, Intergovernmental Panel on Climate Change (IPCC).

MATLAB, 2020. Version 9.8 (R2020a).

McKenzie, G., Hegarty, M., Barrett, T., Goodchild, M., 2016. Assessing the effectiveness of different visualizations for judgments of positional uncertainty. Int. J. Geogr. Inf. Sci. 30, 221–239. https://doi.org/10.1080/13658816.2015.1082566

McMahon, R., Stauffacher, M., Knutti, R., 2015. The unseen uncertainties in climate change: Reviewing comprehension of an IPCC scenario graph. Clim. Change 133, 141–154. https://doi.org/10.1007/s10584-015-1473-4

Millet, B., Cairo, A., Majumdar, S.J., Diaz, C., Evans, S.D., Broad, K., 2020. Beautiful visualizations slain by ugly facts: Redesigning the National Hurricane Center's "cone of uncertainty" map. OSF Prepr. https://doi.org/10.31219/osf.io/wzk8p

Miran, S.M., Ling, C., Gerard, A., Rothfusz, L., 2019. Effect of providing the uncertainty information about a tornado occurrence on the weather recipients' cognition and protective action: Probabilistic hazard information versus deterministic warnings. Risk Anal. 39, 1533–1545. https://doi.org/10.1111/risa.13289

Montello, D.R., Fabrikant, S.I., Davies, C., 2018. Cognitive perspectives on cartography and other geographic information visualizations, in: Montello, D.R. (Ed.), Handbook of Behavioral and Cognitive Geography. Edward Elgar Publishing Limited, pp. 177–196. https://doi.org/10.4337/9781784717544.00018

Morrison, J.L., 1974. A theoretical framework for cartographic generalization with the emphasis on the process of symbolization. Int. Yearb. Cartogr. 14, 115–127.

Nadav-Greenberg, L., Joslyn, S.L., 2009. Uncertainty forecasts improve decision making among nonexperts. J. Cogn. Eng. Decis. Mak. 3, 209–227. https://doi.org/10.1518/155534309X474460

Oberingenieurkreis I, 2014. Vorprojekt Rybibach Lauterbrunnen: Massnahmen für Hochwasserschutz, Tiefbauamt des Kantons Bern.

Olsen, A., 2012. The Tobii I-VT fixation filter: Algorithm description, Tobii Technology.

Olson, M.J., Budescu, D. V., 1997. Patterns of preference for numerical and verbal probabilities. J. Behav. Decis. Mak. 10, 117–131. https://doi.org/10.1002/(sici)1099-0771(199706)10:2<117::aid-bdm251>3.3.co;2-z

Ooms, K., Skarlatidou, A., 2018. Usability engineering and evaluation, in: Wilson, J.P. (Ed.), The Geographic Information Science & Technology Body of Knowledge (1st Quarter 2018 Edition). https://doi.org/10.22224/gistbok/2018.1.9

Padilla, L.M., Creem-Regehr, S.H., Hegarty, M., Stefanucci, J.K., 2018. Decision making with visualizations: A cognitive framework across disciplines. Cogn. Res. Princ. Implic. 3, 1–25. https://doi.org/10.1186/s41235-018-0120-9

Padilla, L.M., Hansen, G., Ruginski, I.T., Kramer, H.S., Thompson, W.B., Creem-Regehr, S.H., 2015. The influence of different graphical displays on nonexpert decision making under uncertainty. J. Exp. Psychol. Appl. 21, 37–46. https://doi.org/10.1037/xap0000037

Padilla, L.M., Kay, M., Hullman, J., 2021a. Uncertainty visualization, in: Balakrishnan, N., Colton, T., Everitt, B., Piegorsch, W., Ruggeri, F., Teugels, J.L. (Eds.), Wiley StatsRef: Statistics Reference Online. pp. 1–18. https://doi.org/10.1002/9781118445112.stat08296

Padilla, L.M., Powell, M., Kay, M., Hullman, J., 2021b. Uncertain about uncertainty: How qualitative expressions of forecaster confidence impact decision-making with uncertainty visualizations. Front. Psychol. 11, 1–23. https://doi.org/10.3389/fpsyg.2020.579267

Padilla, L.M., Ruginski, I.T., Creem-Regehr, S.H., 2017. Effects of ensemble and summary displays on interpretations of geospatial uncertainty data. Cogn. Res. Princ. Implic. 2, 1–16. https://doi.org/10.1186/s41235-017-0076-1

Pang, A., 2008. Visualizing uncertainty in natural hazards, in: Bostrom, A., French, S., Gottlieb, S. (Eds.), Risk Assessment, Modeling and Decision Support. Risk, Governance and Society, Vol. 14. Springer, Berlin, Heidelberg, pp. 261–294. https://doi.org/10.1007/978-3-540-71158-2_12

Pang, A., 2001. Visualizing uncertainty in geo-spatial data. Proc. Work. Intersect. between Geospatial Inf. Inf. Technol. 1–14.

Pang, A.T., Wittenbrink, C.M., Lodha, S.K., 1997. Approaches to uncertainty visualization. Vis. Comput. 13, 370–390. https://doi.org/10.1007/s003710050111

Pappenberger, F., Beven, K.J., 2006. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. Water Resour. Res. 42, 1–8. https://doi.org/10.1029/2005WR004820

Peterson, M.P., 1994. Cognitive issues in cartographic visualization, in: MacEachren, A.M., Taylor, D.R.F. (Eds.), Modern Cartography Series. Academic Press, pp. 27–43. https://doi.org/10.1016/B978-0-08-042415-6.50010-7

PlanGraphics Inc., 1996. The Spatial Data Transfer Standard: Guide for technical managers, Federal Systems Management and Integration Center (FEDSIM), U.S. Geological Survey (USGS).

Potter, K., Rosen, P., Johnson, C.R., 2012. From quantification to visualization: A taxonomy of uncertainty visualization approaches, in: Dienstrey, A.M., Boisvert, R.F. (Eds.), IFIP Advances in Information and Communication Technology. Springer, Berlin, Heidelberg, pp. 226–249. https://doi.org/10.1007/978-3-642-32677-6_15

QGIS.org, 2020. QGIS Desktop 3.16.4 - Hannover with GRASS 7.8.5 - Geographic Information System, QGIS Association.

Quispel, A., Maes, A., 2014. Would you prefer pie or cupcakes? Preferences for data visualization designs of professionals and laypeople in graphic design. J. Vis. Lang. Comput. 25, 107–116. https://doi.org/10.1016/j.jvlc.2013.11.007

Raetzo, H., Lateltin, O., Bollinger, D., Tripet, J.P., 2002. Hazard assessment in Switzerland - Codes of practice for mass movements. Bull. Eng. Geol. Environ. 61, 263–268. https://doi.org/10.1007/s10064-002-0163-4

Ramsey, M.H., 2009. Uncertainty in the assessment of hazard, exposure and risk. Environ. Geochem. Health 31, 205–217. https://doi.org/10.1007/s10653-008-9211-8

Retchless, D.P., Brewer, C.A., 2016. Guidance for representing uncertainty on global temperature change maps. Int. J. Climatol. 36, 1143–1159. https://doi.org/10.1002/joc.4408

Rogowitz, B.E., Treinish, L.A., Bryson, S., 1996. How not to lie with visualization. Comput. Phys. 10, 268–273. https://doi.org/10.1063/1.4822401

Rosenholtz, R., Li, Y., Nakano, L., 2007. Measuring visual clutter. J. Vis. 7, 1–22. https://doi.org/10.1167/7.2.17

Roth, R.E., 2017. Visual variables, in: Richardson, D., Castree, N., Goodchild, M.F., Kobayashi, A., Liu, W., Marston, R.A. (Eds.), International Encyclopedia of Geography: People, the Earth, Environment and Technology. John Wiley & Sons, Ltd, pp. 1–11. https://doi.org/10.1002/9781118786352.wbieg0761

Roth, R.E., 2009. A qualitative approach to understanding the role of geographic information uncertainty during decision making. Cartogr. Geogr. Inf. Sci. 36, 315–330. https://doi.org/10.1559/152304009789786326

RStudio Team, 2020. RStudio: Integrated Development Environment for R (No. 1.3.1073), RStudio, PBC. Boston, MA.

Ruginski, I.T., Boone, A.P., Padilla, L.M., Liu, L., Heydari, N., Kramer, H.S., Hegarty, M., Thompson, W.B., House, D.H., Creem-Regehr, S.H., 2016. Non-expert interpretations of hurricane forecast uncertainty visualizations. Spat. Cogn. Comput. 16, 154–172. https://doi.org/10.1080/13875868.2015.1137577

Salvucci, D.D., Goldberg, J.H., 2000. Identifying fixations and saccades in eye-tracking protocols, in: Proceedings of the 2000 Symposium on Eye Tracking Research and Applications (ETRA '00). ACM Press, New York, pp. 71–78. https://doi.org/10.1145/355017.355028

Scheidl, C., Rickenmann, D., Chiari, M., 2008. The use of airborne LiDAR data for the analysis of debris flow events in Switzerland. Nat. Hazards Earth Syst. Sci. 8, 1113–1127. https://doi.org/10.5194/nhess-8-1113-2008

Schraml, K., Thomschitz, B., McArdell, B.W., Graf, C., Kaitna, R., 2015. Modeling debris-flow runout patterns on two alpine fans with different dynamic simulation models. Nat. Hazards Earth Syst. Sci. 15, 1483–1492. https://doi.org/10.5194/nhess-15-1483-2015

Schweizer, D.M., Goodchild, M.F., 1992. Data quality and choropleth maps: An experiment with the use of color, in: GIS/LIS '92 Annual Conference and Exposition. San Jose, California, pp. 686–699.

SLF/WSL, 2017a. RAMMS::DEBRIS FLOW (Version 1.7.20 64-bit), Institut für Schnee- und Lawinenforschung (SLF), Forschungsanstalt für Wald, Schnee und Landschaft (WSL). Davos, Switzerland.

SLF/WSL, 2017b. RAMMS::DEBIRS FLOW User Manual v1.7.0, Institut für Schnee- und Lawinenforschung (SLF), Forschungsanstalt für Wald, Schnee und Landschaft (WSL).

Slocum, T.A., Cliburn, D.C., Feddema, J.J., Miller, J.R., 2003. Evaluating the usability of a tool for visualizing the uncertainty of the future global water balance. Cartogr. Geogr. Inf. Sci. 30, 299–317. https://doi.org/10.1559/152304003322606210

Slovic, P., Fischhoff, B., Lichtenstein, S., 1982. Facts versus fears: Understanding perceived risk, in: Kahnemann, D., Slovic, Paul, Tversky, A. (Eds.), Judgement under Uncertainty: Heuristics and Biases. Cambridge University Press, pp. 463–489. https://doi.org/10.1017/CBO9780511809477

Stanovich, K.E., West, R.F., 2000. Individual differences in reasoning: Implications for the rationality debate? Behav. Brain Sci. 23, 645–726. https://doi.org/10.1017/S0140525X00003435

Sterba, Z., Blaha, J.D., 2015. Influence of colour on the perception of coherent motion, in: Yaguchi, H., Okajima, K., Ishida, T., Araki, K., Doi, M., Manabe, Y. (Eds.), AIC 2015 - Color and Image. Tokyo, Japan, pp. 533–538. https://doi.org/10.1038/348328a0

Stoffel, M., Huggel, C., 2012. Effects of climate change on mass movements in mountain environments. Prog. Phys. Geogr. 36, 421–439. https://doi.org/10.1177/0309133312441010

Stoffel, M., Mendlik, T., Schneuwly-Bollschweiler, M., Gobiet, A., 2014. Possible impacts of climate change on debris-flow activity in the Swiss Alps. Clim. Change 122, 141–155. https://doi.org/10.1007/s10584-013-0993-z

Swienty, O., Reichenbacher, T., Reppermund, S., Zihl, J., 2008. The role of relevance and cognition in attention-guiding geovisualisation. Cartogr. J. 45, 227–238. https://doi.org/10.1179/000870408X311422

swisstopo, 2019. swissALTI3D: Das hoch aufgelöste Terrainmodell der Schweiz. Wabern.

Tak, S., Toet, A., Van Erp, J., 2014. The perception of visual uncertainty representation by non-experts. IEEE Trans. Vis. Comput. Graph. 20, 935–943. https://doi.org/10.1109/TVCG.2013.247

Thompson, M., Warmink, J.J., 2016. Natural hazard modeling and uncertainty analysis, in: Riley, K., Webley, P., Thompson, M. (Eds.), Natural Hazard Uncertainty Assessment: Modeling and Decision Support. American Geophysical Union, pp. 11–19. https://doi.org/10.1002/9781119028116.ch2

Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M., Pavel, M., 2005. A typology for visualizing uncertainty, in: Proceedings Volume 5669, Visualization and Data Analysis 2005. San Jose, CA, United States. https://doi.org/10.1117/12.587254

Tobii AB, 2022a. Understanding Pro Lab's eye tracking metrics [WWW Document]. URL https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/data/understanding-Pro-Labs-eye-tracking-metrics/ (accessed 10.1.22).

Tobii AB, 2022b. Eye tracking metrics [WWW Document]. URL https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/interpret/eye-tracking-metrics/ (accessed 10.1.22).

Tobii AB, 2021. Tobii Pro Lab (Version 1.171.34906). Tobii Pro AB, Danderyd, Sweden.

Tobii Pro AB, 2014. Tobii Pro Lab User Manual (Version 1.171). Danderyd, Sweden.

Tobii Technology, 2012. Determining the Tobii I-VT Fixation Filter's default values: Method description and results discussion, Tobii White Paper.

Trau, J., Hurni, L., 2007. Possibilities of incorporating and visualising uncertainty in natural hazard prediction maps.

Tversky, A., Kahnemann, D., 1982. Availability: A heuristic for judging frequency and probability, in: Kahnemann, D., Slovic, P., Tversky, A. (Eds.), Judgement under Uncertainty: Heuristics and Biases. Cambridge University Press, pp. 163–178. https://doi.org/10.1017/CBO9780511809477

Tversky, A., Kahnemann, D., 1974. Judgment under uncertainty: Heuristics and biases. Science (80-.). 185, 1124–1131. https://doi.org/10.1126/science.185.4157.1124

USGS, 1997. Spatial Data Transfer Standard (SDTS): Logical specifications, U.S. Geological Survey (USGS). Virginia.

Uusitalo, L., Lehikoinen, A., Helle, I., Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. Environ. Model. Softw. 63, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017

Van Der Bles, A.M., Van Der Linden, S., Freeman, A.L.J., Mitchell, J., Galvao, A.B., Zaval, L., Spiegelhalter, D.J., 2019. Communicating uncertainty about facts, numbers and science. R. Soc. Open Sci. 6, 1–42. https://doi.org/10.1098/rsos.181870

Wallsten, T.S., Budescu, D. V., 1995. A review of human linguistic probability processing: General principles and empirical evidence. Knowl. Eng. Rev. 10, 43–62. https://doi.org/10.1017/S0269888900007256

Walser, M., McArdell, B., Huggel, C., Graf, C., 2014. Inwiefern eignen sich numerische Modellsimulationen für die Gefahrenkartierung von Murgängen - ein Vergleich. Agenda FAN 21–23.

Walther, D., Koch, C., 2006. Modeling attention to salient proto-objects. Neural Networks 19, 1395–1407. https://doi.org/10.1016/j.neunet.2006.10.001

Ware, C., 2013. Color, in: Ware, C. (Ed.), Information Visualization: Perception for Design. pp. 95–138. https://doi.org/10.1016/B978-0-12-381464-7.00004-1

Wilkening, J., Fabrikant, S.I., 2011. How do decision time and realism affect map-based decision making?, in: Egenhofer, M., Giudice, N., Moratz, R., Worboys, M. (Eds.), Spatial Information Theory. COSIT 2011. Lecture Notes in Computer Science, Vol 6899. Springer, Berlin, Heidelberg, pp. 1–19. https://doi.org/10.1007/978-3-642-23196-4_1

Windschitl, P.D., Weber, E.U., 1999. The interpretation of "likely" depends on the context, but "70%" is 70%—right? The influence of associative processes on perceived certainty. J. Exp. Psychol. Learn. Mem. Cogn. 25, 1514–1533. https://doi.org/10.1037//0278-7393.25.6.1514

Witt, J.K., Clegg, B.A., 2021. Dynamic ensemble visualizations to support understanding for uncertain trajectories. J. Exp. Psychol. Appl. 1–17. https://doi.org/10.1037/xap0000370

Wittenbrink, C.M., Pang, A.T., Lodha, S.K., 1996. Glyphs for visualizing uncertainty in vector fields. IEEE Trans. Vis. Comput. Graph. 2, 266–279. https://doi.org/10.1109/2945.537309

Wobbrock, J.O., Findlater, L., Gergle, D., Higgins, J.J., 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11). Association for Computing Machinery, Vancouver, BC, Canada, pp. 143–146. https://doi.org/10.1145/1978942.1978963

Wolfe, J.M., Horowitz, T.S., 2004. What attributes guide the deployment of visual attention and how do they do it? Nat. Rev. Neurosci. 5, 1–7. https://doi.org/10.1038/nrn1411

Zhang, J., Goodchild, M., 2002. Uncertainty in geographical information, 1st ed. Taylor & Francis, London. https://doi.org/10.1201/b12624

Zimmermann, M., Pozzi, A., Stoessel, F., 2005. Hazard maps: The Swiss system and its application applied abroad. Vadem. Swiss Agency Dev. Coop. SDC 1–34.

Zuk, T., Carpendale, S., 2006. Theoretical analysis of uncertainty visualizations. Vis. Data Anal. 6060. https://doi.org/10.1117/12.643631

8. Appendix

A. Transcript of Expert Interview E1

Interviewer:	Interviewee:	Role:	Date:
Jana Bracher	E1	Scientific community	09.08.2021

This interview was conducted in English. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability. The transcript was later edited by the interviewee.

JB	*First of all, thank you very much for participating. And like you said you are okay with the recording.*
E1	Yes.
JB	*Okay first of all, please just describe shortly some common characteristics of a debris flow.*
E1	The process itself is intermediate between landslides and floods. So, they tend to travel in channels, like in floods, but they often look like a landslide, sometimes they even appear to be somewhat dry looking as they flow down a channel. The fronts often look dry, and a big difference is that a flood, when it travels down a channel, it carries, it drags the boulders with it. So, the big particles maybe travel a factor of 10 slower than the water itself. Whereas in a debris flow everything moves at the same speed. And that's one of the things that makes them so hazardous. So, you can think of them as an ultra-mobile kind of landslide. In Switzerland they are treated with the water hazards, which is interesting. In, I think in most other countries in the world they are treated as a landslide hazard.
JB	*What are the main causes of debris flows?*
E1	Well, you have to have three conditions to get debris flows. You have to have steep slopes and you have to have a lot of gravitational force acting on the sediment. You have to have a lot of sediment and you have to have a concentrated source of water. So, rainfall is number one in Switzerland, but also glacial outburst floods or lake outburst floods are causes of debris flows. But most of them are caused by very intense rainfall.
JB	*And maybe specific to Switzerland when looking at the whole variety of natural hazards, how big would you judge the importance of debris flows in that context?*
E1	I think hail from a damage perspective, hail is number one, but we don't study hail here but and then the landslides are the second most important probably, and floods. Floods may cause more damage than landslides, I am not sure, a group at WLS has worked up some statistics, I can send you that most recent summary from them.
JB	*And what do you think, because you said that debris flows are very hazardous, what are the most common damages that occur actually?*
E1	It's mostly damaged infrastructure and property damage it's like railway lines, roadways, bridges occasionally houses. And then fortunately people aren't directly affected that much thanks, largely, to the land use planning that we have in Switzerland. I think there is roughly a fatality every two years on average due to debris flows. It's hard to separate. It's hard to separate material damage from the human damage. But I think landslides kill more people than debris flows. Again, I will send you the specific paper from this. And

	flooding is also hazardous more in terms of property than human damage. But occasionally people are also trapped in cellars like they were recently in Germany. But I am not an expert on flooding.
JB	*And then what are some of the most important measures that can be taken to prevent these damages?*
E1	Land use planning is number one and Switzerland is really at the forefront. And that is the hazard maps that every community is required to have. They're the basis for land use planning. And so, if you own a house in the red zone, you will not be able to take on a loan of a 'Hypothek' on the house for example. So, the land use planning is quite good at keeping people away from hazardous areas. Then the emergency planning that every community is doing also helps a great deal. So, when a storm event happens, the hazard managers know exactly what to do. They've planned in advance what they should do. I think that helps a lot.
JB	*And in terms of preventing the actual event, that is pretty much not possible right?*
E1	There are some things that can be done is to drain hazardous lakes, for example in the Glaciers above Lenk, they have drilled tunnels to drain the lakes to prevent the hazardous extreme floods that could turn into debris flows. In terms of debris flows themselves, there's some measures that are done to stabilize the land surface to prevent the initial landslides, so that helps. Retention basins, basins of course they don't stop a debris flow, but they prevent it from reaching a village, but those are very expensive to maintain.
JB	*And then in case of a debris flow event, are those events usually mapped and visualized? Or how is the data recorded?*
E1	Good question, the Swiss for federal office of the environment, the BAFU, they have a form that hazard workers are supposed to fill out for damage-causing events, called StorMe. And that is supposed to have a basic characterization of the event: how large, what's affected, what's damaged, where it started, where it ended, as well as estimates of the precision. And if there is a fatality, then things are generally mapped extremely well, because they want to understand exactly what happened. Every time there is a large event, a very large event that was unexpected or larger, something on the order of a 100-year return period, the community will hire a consultant, who will update the hazard map.
JB	*And then what is usually visualized? Is it just the spatial extent of the event or flow heights or intensities?*
E1	Yes, intensity is really difficult because that involves both flow height and velocity. And velocity is really difficult to estimate. But they usually involve a spatial extent and then also, the depth. The deposit depth, which is part of the mapping criteria for hazard maps, and then also for the maximum of the flow depth, so they'll measure mud-line heights on trees and houses.
JB	*Okay, and when those maps are produced what are the… so in cartography you work with certain visual variables. So, like colour hue or texture. Do you know whether there are some specific visual variables that are used in those maps?*
E1	Yes, the intensity there is a standard colour scheme that is encouraged by the BAFU. It involves shades of green and yellow. Dark green is most intensive and yellow is the least intensive. (…)
JB	*Do you think maps are an important medium when dealing with debris flows or natural hazards in general, whether it's in decision making or in research?*
E1	I think they're absolutely critical. For visualization, I think a lot of people, especially the laypeople, not the people involved in technical stuff, people who live there. They look at the map, they see where their house is, they see all the houses and that is a very important part of communicating the spatial extent of the hazard, I think they are critical. Otherwise, people don't have a sense of where the hazard is.
JB	*Yes, I think the visual form is kind of, a very straight forward way to show the hazard.*
E1	I completely agree. It's a very good question but I think without the maps there would be a great deal of confusion. Of course, we know there are many people who can't read the map. But I think if they can identify

	their house or other objects like a church, a sport field or something then they get a better sense. Good question.
JB	*How are future debris flows predicted? It is even possible? Or are for example past extents used for predictions?*
E1	Oh, prediction is, is something we'd all like to do, it's a great deal of effort going into predicting debris flow activity and I'd say the hardest to predict is how large a debris flow will be and exactly where it will start. Several factors, let me back up. Predicting where an initial landslide or debris flow will start is very difficult. Predicting when they will occur is a little bit easier because they tend to happen during very intense rainstorms. Where they will flow is somewhat straightforward because they follow topography quite strongly. But how big they will be is very difficult to estimate in advance. The larger a flow is the farther it will travel. And… everyone would love to be able to predict how big a debris flow will be. Many of us spend a lot, far too much time working on that but it's very difficult.
JB	*So, you would say the spatial extent is the hardest, also when generating hazard maps in those regions?*
E1	Yes, I would say that's why people tend to work with scenarios. So, they have a scenario that they'll say for a largest expected event, and they'll look at past historical events and they'll say: okay the event we had 100 years ago went this far, let's just make an assumption that this is close to a 100-year return period event. There are some hydrological tools they can use to estimate the return period of at least the rainstorm.
JB	*And when we apply these difficulties to the uncertainties. Where do you think are the largest uncertainties? Is it again more of the spatial uncertain or temporal or also because we speak about positional uncertainty, attribute uncertainty and temporal ones.*
E1	That's a very good question. I think, temporal is probably a little bit easier, because during periods of dry weather people simply don't expect flows. Of course, there are the occasional events that do happen, like in 2001 in Täsch there was a proglacial lake that failed and created a large debris flow on a sunny day. But in general, I'd say the spatial is more difficult to predict. Simply because small flows will remain in the channel, but the larger they are the more likely they will leave the channel and go to places that people don't expect.
JB	*Looking at, you said you always look at scenarios or do scenario, where do you think are those uncertainty introduced? Are they kind of a problem of phenomenon itself, or are uncertainties also introduced within the simulations?*
E1	Another great question. You will learn about that in a couple of weeks in great detail. Number one is volume, event volume and how that volume fails. Will it come as one huge surge, that's probably worse than if it comes with two smaller surges. Probably, not always. There is some uncertainty… we use models to predict how fast they go and how far they go and there's some coefficients you need for these models. And there's some uncertainties associated with those as well.
JB	*Okay so they are introduced in the process of creating simulations as well?*
E1	That's right. And the flow can change as it flows down a channel. It can entrain water, it can entrain dry sediment and that will change the so-called frictional properties and that changes the coefficients that we use in the model. But again, I think the uncertainty in volume is number one.
JB	*So, is that also the aspect of uncertainty that you as a researcher are most interested in or is the temporal something that is also interesting to you?*
E1	Good question. I work, I spend time trying to come up with better ways to estimate the volume of debris flows at the I. observation station. Because we don't have a basin that traps everything. So, I have to use the speed of the flow as it's going past my station along with some assumptions about the rheology (the material properties of the flow) to try to come up with better ways of estimating the volume. I spend most of my time doing research about the flow properties.
JB	*Okay, and what are the largest uncertainties regarding this volume in your work?*

E1	So you go into a catchment, say somewhere above L.. And you look at the landscape and you try to identify places where a lot of sediment has accumulated. And then you try to estimate how much sediment there is available and then the harder part, is to estimate how much sediment can be mobilized by a debris flow. And then another hard part is estimating how much is going to fail when it rains. Because not all the sediment will come at once.
JB	*Okay, and when depictions of debris flow are generated, are uncertainties usually depicted? Or maybe just mentioned in descriptions?*
E1	I am most familiar with hazard maps. And hazard maps don't have much in the way of uncertainty in them. They look quite deterministic and finished. There are of course estimates of uncertainty, that go into them.
JB	*And are those debris flow visualizations used among experts and for experts of also for the public?*
E1	Ultimately, they're for the public. In, but they're used for land use planning. Many of the Cantons now have hazard maps available online in a data portal that everybody can access. But many people don't know how to read or interpret them.
JB	*And have you personally ever experienced difficulties when communicating uncertainties to maybe those practitioners or decision makers? Or is that something that is not discussed?*
E1	I haven't done that much applied work. So, no.
JB	*And do you expect that including uncertainty in those visualizations would help the decision makers actually or would they just be confused by all the mathematical information for example?*
E1	Ah, that's a good point. My impression is that the practitioners already approach the problem with a sense of uncertainty. But the hazard map shows sharp boundaries with no uncertainty. I think it would help but it would require training though. I think it would help to give practitioners a better sense of the uncertainties. I think that would make it more objective in the end.
JB	*And if you were to visualize uncertainty for example in debris flow simulations, how would you do it as an expert, coming from your perspective?*
E1	I would prefer to work within the existing successful mapping framework and instead try to show some additional uncertainty, for example between the red and blue shaded areas. However, a major problem I have not yet thought through is that the consequences are probably not as flexible. Should a bank lend money to a homeowner to modify their house if the house was in a location that was somewhere between the blue and red (construction prohibition) zones?
JB	*present maps of hurricane cone and my sketch* *And then now, like I wrote you in my email, my inspiration is basically coming from the hurricane cone of uncertainty, that is used in the US. So, I would quickly like to show you how this cone of uncertainty looks. I am not sure whether you are familiar with it?*
E1	I have seen the maps, but I haven't actually looked into what creates the cone of uncertainty.
JB	*The way it's displayed it has a point in time that is the now basically in the forecast, it includes a line of the most likely path that the hurricane is going to take and then it also includes I think the 60% probability where the path could go. So, it includes like 60% of uncertainty and the problem is basically that a lot of people misinterpret the concept thinking that the hurricane will get bigger by the time, which is not ethe case, it just shows a variety of the paths.*
E1	Yes, I see the problem, alright.
JB	*And now I would like to show you a little sketch that I made trying to apply this concept to a debris flow. So here I chose the Rotlauwi place in Guttannen and this is basically one of my concepts, a very beginning one, where we are using transparency actually to depict the uncertainty, the spatial uncertainty in debris flows. So, showing places which the debris flow is less likely to reach more transparent than those that are probably going to be*

	affected. Now, do you think this broad concept of the uncertainty cone of the hurricanes could be something that is applicable to debris flows?
E1	It could be, there's some differences of course. Hurricanes are travelling, at least on this map that you showed they are travelling over flat terrain, debris flows tend to follow topography quite strongly. So, to the extent that a debris flow will deviate from a flow path is related to the size of a flow. So, a large flow will become larger and then flow overbank and go to other areas. But there is some additional uncertainty that for example a small landslide might occur and that might block the channel and cause the debris flow to leave the channel and go somewhere where it otherwise wouldn't. For example, at a bridge crossing: people might just assume the bridge will be destroyed but maybe the bridge will survive and that will cause the flow to build up and then it will switch and go into an area that was not expected. So, I think it could be done, I think you'd probably want to use either some topographic index to show the flow lines on the topography maybe some kind of D8 algorithm.
JB	*Okay maybe coming back to the cone of uncertainty, here, would you agree that it is more of a temporal and spatial uncertainty, while depicting uncertainty in debris flow would need to be more spatial because the temporal aspect is such a short timeline.*
E1	I would agree with that, I think that is a very fair way to put it.
JB	*Now coming back who is going to use those maps, so the decision makers. What are typical decisions, people have to take with the help with depictions of debris flow hazards?*
E1	I have never worked at this level, but I imagine that decision makers have to decide if a parcel of land can be used for new houses or other constructions such as a school or a hospital.
JB	*Yes, what are some typical decisions that are made with the help of the hazard maps or more specifically debris flow simulations?*
E1	Yes, starting with the hazard maps decisions have to be made: it's in the yellow zone, what is the actual hazard? Could something be done to provide additional protection? I think those are quite important things.
JB	*One last topic I would like to talk about is the influence of climate change on debris flows, So, how are the characteristics of debris flows in Switzerland expected to evolve due to climate change?*
E1	Okay, you'll be meeting with E2 I guess?
JB	*Yes.*
E1	Okay, great. He'll tell you more about this (*laughs*). But specifically, we've really only worked at one location. The debris flow season traditionally is between May and October, now is longer or will become longer in the future. So, the entire year is in principle possible for lower elevations anyway.
JB	*What are the uncertainties that are tied to those debris flow predictions due to climate change? Maybe very broadly.*
E1	(...)
JB	*What could be the influence of thawing of permafrost have on debris flows due to climate change?*
E1	Main factors include increased sediment availability and production. More sediment could be available for debris flows if the active layer frozen less often than in the past, and increased numbers of freeze-thaw cycles at higher elevations could cause more sediment to be released from mountainsides, thereby increasing the amount of sediment potentially available for debris flows.
JB	*Okay, these were all the questions I prepared, is there any aspect of debris flows and their uncertainties that we have not touched on and that you would like to mention?*
E1	One aspect: we tend to put things into bins. We say this a debris flow, this is a landslide, this is a flood. Nature tends not to follow our bins. We get events that show aspects of all these processes. And I think, that's a major challenge, and one expert's landslide is another person's debris flow. And one may start as a debris

	flow, but the damage is caused by a flood that happened further downstream and no longer looks like a debris flow. We know what to do with a debris flow, we know best how to manage a flood and a landslide, but we have the transitional events, events that are in between landslides and debris flows or in between floods and debris flows.
JB	*Okay well, thank you very much, it has been really interesting.*
E1	Well, thank you.

B. Transcript of Expert Interview E2

Interviewer:	**Interviewee:**	**Role:**	**Date:**
Jana Bracher	E2	Scientific community	09.08.2021

This interview was conducted in English. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability.

JB	*Thank you for doing this interview with me. Like you said you are okay with the recording?*
E2	Yes.
JB	*Okay, so let's get started. Just some basic questions about debris flow, maybe even though you are new to debris flows.* *Please describe the common characteristics of a debris flow.*
E2	So usually they occur in places where it's rather steep and they can be triggered by rainfall or earthquakes or even snow melt and here in Switzerland, the Alps, I would we consider more debris flows triggered by rainfall and snowmelt than earthquakes. And they occur in channels, it's a channelized process usually, and high sediment concentrations, 40% or more. And they have another flow behaviour than water.
JB	*In the context of all natural hazards in Switzerland, how high would you rank the importance of debris flows?*
E2	It's definitely smaller than floods. Often it's also not so easy to… they occur at the same time sometimes and it's not so easy to separate. Floods and hail are actually the most important ones in Switzerland and then followed by landslides I would say and then maybe debris flows.
JB	*What are the most common damages resulting from debris flows?*
E2	Well, I don't know any statistics, there do exist some probably. But I would say streets that are closed, railways that are closed, sometimes houses that are impacted.
JB	*What are some of the most important measures that can be taken to prevent these damages?*
E2	Construction measures or early warning.
JB	*From your experience, how are (past) debris flow events and model simulations commonly mapped / visualized? And are they even mapped?*
E2	I mean they are probably for the hazard maps. But I don't have that much experience with that. And I don't know of any database where all debris flows in Switzerland would be mapped.
JB	*What is the importance of maps as a way of presenting information in the context of debris flows / natural hazards? Do you think it's a good tool to help understand processes or understand projections of debris flows?*

E2	I guess, they are always helpful, so yes I think so, that's definitely a good thing to do, especially since, it's good to understand where they can occur.
JB	*In the context of predicting future debris flows. How are future debris flows predicted? Are for example past event extents used as a standpoint to predict future ones?*
E2	Yes, definitely most prediction models are empirical or calibrated with past events. So we are very dependent on past events for predictions. And… there were two questions.
JB	*Yes, whether past events are used, so you already answered both in once. What are the greatest difficulties when predicting future debris flows?*
E2	There are many factors playing a role and each of these factors have their uncertainties. And uncertainties add up of course, so… It's difficult to see in which direction it could go.
JB	*Within the context of visualizing uncertainty we distinguish between for example positional uncertainty, attribute uncertainty and also temporal uncertainty. Where do you think are the biggest uncertainties when predicting debris flows?*
E2	So I am not aware of these terms. I mean for sure there is some temporal uncertainty because, well you never know when exactly. And what were the other two?
JB	*Positional uncertainty, so kind of spatial uncertainty and also attribute uncertainty as for example, with the debris flow for example the flow height, so not knowing how high the flow height will be.*
E2	(…) There is also a different importance, I guess in this kind of uncertainty, but positional uncertainty is difficult because… well you don't really, you know, you can estimate for example where permafrost will thaw, but you cannot… but there are many other factors involved and if there is really a potential to trigger a debris flow. So positional uncertainties are very high I would say. And, yeah and how high it will be at some place is basically impossible, no not impossible, but it's very difficult to say.
JB	*And you in your own research, what aspects of uncertainty are you most interested in? (spatial, temporal attribute)*
E2	So, we mainly looked at uncertainty in the climate forcing. So, one thing is: how much warmer will it be or how will the climate change, but the other thing is also that there is an internal uncertainty in the climate, because you can have… you can have dry, even in stable climate, you can have very dry or very wet years. I mean the climate has not changed in the last two or three years but we had a drought in 2018 and we had a very wet 2020. It's important to, to get an idea of this internal climate uncertainty.
JB	*Okay so not even the uncertainty directly regarding the debris flow, but kind of on a higher level, those uncertainties.*
E2	Yes, this would be from, yeah, considers more the triggering of debris flows.
JB	*Coming back to the spatial uncertainties. What are the largest uncertainties when it comes to estimating the spatial extent for future debris flows?*
E2	The largest uncertainties for the spatial extent… (…) I guess you, well, there needs to be sediment somehow and there are different processes involved in sediment generation. And even getting an idea of which are the driving forces in sediment generation can be very tricky.
JB	*To your knowledge, is uncertainty something that is visualized for example in debris flow simulations?*
E2	(…) Often no (*laughs*). I mean it requires that you do some kind of, or at least in the quantitative context, it requires you do some kind of sensitivity analysis. And in science yes, often done, but in practice or I am at least not aware in practice.
JB	*And are those debris flow visualizations produced more for experts, so to use in the scientific community, or also for the public?*

E2	Mostly in the scientific community, I would say, at least for me it has been like this so far. There are some interactions, but…
JB	*Have you personally ever had some issues or difficulties when communicating uncertainties in debris flow predictions whether it's within the scientific community or towards the public or decision makers for example?*
E2	In the scientific context no, I mean of course you have to explain but if you talk with scientists they will mostly understand. And I have not had that much exchange with other parties, they have small questions, so I cannot answer this.
JB	*And do you think including uncertainty visualization would actually help people reading the maps who have to make decisions?*
E2	I think yes, it's a bit hard to image how this visualization would really look like but in general visualizations are helpful, so…
JB	*If you were to create such a visualisation to depict how debris flows are going to evolve for example in the prospect of climate change, how would you go about this?*
E2	In Switzerland for example?
JB	*Yes.*
E2	I think I would start from rainfall and maybe visualizing somehow how rainfall thresholds could, or rainfall above a specific threshold could change, so how extreme events could change. And then you get the spatial idea of where do we expect many changes in extreme rainfall. And, so this is more the rainfall forcing part. And then it's interesting to see where are which kinds of sediment availability or sediment generating processes. But I would, I don't know where I would start this.
JB	*Yes, I think it's a difficult task.*
E2	And permafrost is of course always… how will the permafrost change.
JB	present maps of hurricane cone and my own transparency sketch *And then in my thesis, my inspiration for uncertainty visualisation basically came from hurricane cone visualizations. Here is one of these visualizations, I am not sure whether you have seen it before, but basically the concept is that you start at a point in time, so these maps are created very dynamically in advance of an event. So this is kind of the time point now and then it shows the line of the most likely path the hurricane is going to take. And then it also shows this cone which is the 2/3 possibility or uncertainty for the path which is based on past events or past trajectories of hurricanes. And now in my work, like I said, my aim is to apply a similar concept to debris flows. So here is a little sketch that I made, it's not based on real data, in Guttannen of the Rotlauwi. So here it basically shows the concept to create this cone of where a debris flow could occur or what areas could affected by a debris flow. And then also showing of course towards the other edges the uncertainty that this place will be reached is higher, so it's going to be displayed more transparent and the places within the riverbed that are likely to be affected are displayed less transparent.* *Do you think the concept of the hurricane cone is something that is applicable to debris flows actually?*
E2	Yes definitely. This would be… so for the hurricane is this used for short term prediction or is it more like, being aware of…
JB	*Yes, this is more for a short term prediction so basically the purpose of this is that there is going to be a hurricane alert and then these maps are produced and then also broadcasted on TV. And the purpose of debris flow visualization would be more, uhm, as a basis for decision making for more long term planning because of course, yeah, debris flows are more spatial and I think hurricanes are probably more spatial and temporal as well.*
E2	I wonder what is the, what would you say, how does this differ from a hazard map for example?
JB	*Yes, that's the purpose, so basically this (visualization) could also be an input for generating hazard maps also in the end.*

E2	Okay, so, but you have this uncertainty range (points at map legend).
JB	*Yes, kind of the base purpose is to depict the uncertainty that's implicit to the event.*
E2	For sure it's very useful to have something like this, the question is how you generate it. And how you compute this uncertainty. But the concept, I like very much.
JB	*Looking at a map like this and also thinking of the hazard maps as well, what are some common decisions, that practitioners or decision makers need to make with projections of debris flows?*
E2	Sorry could you repeat that?
JB	*What are some common decisions, that decision makers need to make with the help of hazard maps or also with visualizations of debris flow projections?*
E2	Probably how they, yeah exactly, what is behind this? How do they compute these uncertainties? Or how do they generate, where do they get… do rely only, do they rely on models, do they only rely on mapping past events? Or both? And yeah, I am not sure.
JB	*Maybe now we come a little bit more to your topic about debris flows with climate change. How are the characteristics of debris flows in Switzerland expected to change due to climate change?*
E2	I mean this can be very different in any places. But we have looked at the I., which is a very active site in V.. And on the one hand our simulations showed that the hydrology will change in a way that there will actually more debris flows, if there is always enough sediment there will be more debris flows, also, not so much in summer because summers will also be drier. But it's really an elevated activity in spring, in the autumn and a few events even in winter, which don't haven now. And then, well the other factor is, is there sediment? And we think that sediment is generated by frost cracking in I.. But since I. has no permafrost, so there will, we don't expect more sediment availability due to thawing and frost cracking, this activity will actually decrease because you get fluctuations around zero degrees. Because it gets warmer. If you maybe now it fluctuates around zero but if it's three, four, five degrees warmer, it fluctuates around five degrees and not around zero anymore. So you loose this process. And if we consider all this there is the possibility of having less debris flows in I.. But it also clearly shows that, looking at the hydrology there is more potential. And in other places, for example permafrost places, it's clearly a higher risk of having more debris flow.
JB	*So permafrost could have a big influence on debris flow activity at some sites?*
E2	Yes.
JB	*And then I guess the predictions are also very site specific?*
E2	I mean kind of. Yes you have to do it for every site for really quantitative estimates. But, we know that there will be less snow, there will be more, some more intense rainfall. And this is quite consistent over Switzerland at least, even the Alps.
JB	*What are the uncertainties when trying to do those prediction for the future?*
E2	So the first is, we rely on observations and you don't really have this kinds of observations on other places, so it's difficult to calibrate your models. And if you cannot calibrate your models you have higher uncertainties. And then it's really, I would say, in I. we think, it's frost cracking, is the driving process. But even in I., which is a few square kilometres, it's, it's difficult to really say, this is the driving process. So, and it might be some other processes in other places. And identifying these sediment generation processes is very challenging.
JB	*Yes, so the uncertainties are not introduced by the debris flow itself but the whole processes happening in the landscape?*
E2	Yes, I would say. I mean we have not looked… we have more looked at how triggering conditions change. Also in the context of climate change, there might be a heavier rainstorm in July but it might fall on dry ground because there are also more droughts in summer. And, this will impact the runout as well. And that's

	very difficult to say. On really climate change impacts on debris flow runout is, I don't know any studies that do this so far. So we really looked more on the triggering conditions.
JB	*Yes that was already my last question. Maybe from your side, are there any aspects of debris flows and their uncertainties that we have not discussed and that you would like to mention?*
E2	Yes, maybe just adding to the last point that. Because of this high or broad range of climate, what I said earlier already, even in a stable climate the range is quite wide, so it's actually quite difficult to see changes. So even with the results I was saying from our simulations, there is always overlapping of different simulations. So, uncertainties in climate are even very, are large and they are transmitted to the debris flow predictions as well of course. I think that was it.
JB	*Okay, great! Thank you very much for talking to me, it was very interesting.*

C. Transcript of Expert Interview E3

Interviewer:	**Interviewee:**	**Role:**	**Date:**
Jana Bracher	E3	Scientific community	10.08.2021

This interview was conducted in Swiss German and translated to German during the transcription which resulted in the rephrasing of certain statements. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability.

JB	*Wie du gesagt hast, für dich ist es in Ordnung, das Interview aufnehmen.*
E3	Ja.
JB	*Sehr gut, als erstes bitte ich dich die generelle Charakteristik von Murgängen zu beschreiben.* *Please describe the common characteristics of a debris flow.*
E3	Ja, Murgänge sind ein Gemisch aus Sediment und Wasser mit Anteil 50/50 oder auch variabel, manchmal auch mehr Sediment, welches sich bildet bei Niederschlag. Sie können kanalisiert sein oder am Hang, also Hangmuren oder kanalisierte, welche sich schnell bewegen, grosse Abflüsse generieren können, sehr viel grösser als normale Hochwasser und auch einfach sehr viel destruktive Gewalt haben, deshalb sind sie auch sehr gefährlich auch für Menschen.
JB	*Und was sind die Hauptauslöser von Murgängen?* *What are the main causes of debris flows?*
E3	Normalerweise sind es eigentlich starke Regenfälle, also wenn der Boden auch gesättigt ist. Und dann in steilen Gebieten, Teile des Bodens oberflächlich abrutschen können und sich zusammenfinden können und solche Murgänge bilden. Es kann aber manchmal auch sein, zum Beispiel gibt es Seeausbrüche, Geltscherseeausbrüche, welche anschliessend zu Murgängen frühen können. Solche grosse Releases, wenn grosse Wassermassen schnell abfliessen in instabilem Material, kann das auch zu Murgängen führen.
JB	*Im Kontext von allen Naturgefahren in der Schweiz, wie hoch würdest du die Relevanz von Murgängen einschätzen?* *In the context of all natural hazards in Switzerland, how high would you rank the importance of debris flows?*

E3	Also im Mittelland ist es natürlich nicht so relevant, dort gibt es praktisch keine Murgänge. Vielleicht ab und zu solche Hangmuren. Oder so in Hanggebieten schon aber dort ist es nicht so relevant. Aber in den Bergen natürlich schon, in den Voralpen, ab den Voralpen bis zu den Alpen ist es definitiv eine der wichtigsten Gefahren vielleicht nebst Schneelawinen und Hochwasser.
JB	*Welche Art von Schäden entstehen generell durch Murgänge?* *What are the most common damages resulting from debris flows?*
E3	Murgänge sind sehr destruktiv, sie können grosse Sachen… es kommt natürlich darauf an, wie gross der Murgang ist, wie hoch der Abfluss ist. Weil sie auch viel festes Material enthalten, sind sie sehr destruktiv, also könnten locker ein Haus abräumen. Und natürlich, ja Menschen, Menschenleben sind extrem gefährdet, das überlebt man nicht, wenn man da rein gerät.
JB	*Welche Massnahmen können ergriffen werden, um diese Schäden zu verhindern?* *What are some of the most important measures that can be taken to prevent these damages?*
E3	Also das beste ist eigentlich die Raumplanung; nicht zu Bauen in den Gebieten, welche gefährdet oder stark gefährdet sind in den Gefahrenkarten. Das wird eigentlich normalerweise, also wenn es Gebäude hat in der roten Zone, dann werden Massnahmen am Bach oder so ergriffen oder man muss… Es gibt es wenig, dass man umsiedeln muss, zum Beispiel in Bondo oder ein paar andere, da war das ein Thema. Frühwarnsysteme sind sicher ein gutes Mittel, um Menschenleben, oder irgendwelche mobile Gegenstände aus der Gefahrenzone zu bringen.
JB	*Also geht es bei der Massnahmenbekämpfung mehr darum quasi die Folgen zu verhindern, aber das Phänomen selber kann man gar nicht verhindern?* *So those measures are more about preventing the consequences of debris flows and the phenomenon itself basically cannot be prevented?*
E3	Ja, also das Phänomen ist… also ja man macht zum Teil solche Retentionsbecken, sodass sich das Sediment setzen kann. Das ist eigentlich keine schlechte Massnahme. Das sieht man immer wieder solche Sedimentbecken in den Alpen, das geht schon.
JB	*Und wenn es zu einem Ereignis kommt, werden Murgänge normalerweise kartiert oder visualisiert im Nachhinein?* *How are (past) debris flow events commonly mapped / visualized?*
E3	Nein, nicht unbedingt, sie sind eigentlich wirklich vor allem durch die Gefahrenkarten… also durch die Gefahrenkarten sind Murgänge ziemlich gut abgedeckt. Das ist wirklich über die Gefahrenkarten, in denen diese Phänomene abgedeckt sind.
JB	*Welche Parameter werden kartiert, beispielweise die Intensität oder nur der räumliche Bezug?* *What parameters are visualized? E.g., intensity, flow height*
E3	Es ist eine Kombination von Intensität und Wahrscheinlichkeit. Und die Intensität ist eigentlich vor allem die Abflusshöhe oder die Ablagerungshöhe und die Geschwindigkeit.
JB	*In der Kartografie arbeitet man mit visuellen Variablen, beispielsweise die Verwendung von Texturen oder Farben zum kartieren. Weisst du, gibt es irgendwelche Standards, wie man Murgänge oder Gefahrenkarten darstellt?* *What visual variables are used? E.g., colour hue, colour value, texture etc.*
E3	Also bei den Gefahrenkarten gibt es den Standard: rot-blau-gelb. Bei den Prozesskarten oder Modulierung dort kann man, dort hat man mehr Spielraum, dort kann man beispielsweise für die Abflusshöhe oder die Ablagerungshöhe eine Farbpalette nehmen, dann es semitransparent machen. Aber für die Gefahrenkarten selber ist es sehr streng.
JB	*Und bei Modellierungen wird der Murgang selber als Prozess dargestellt?*

E3	Ja.
JB	*Welcher Stellenwert ist Karten zuzuordnen im Umgang mit Murgängen?* *What is the importance of maps as a way of presenting information in the context of debris flows / natural hazards?*
E3	Karten würde ich sagen vor allem, ja, über die Gefahrenkarten. Oder dann als Gefahrenhinweiskarten, dort werden eher die Prozesse etwas grösser… die Skala ist normalerweise 1:50'000, Gefahrenhinweiskarten. Die Gefahrenkarten sind 1:20'000 / 1:10'000. Und bei der Gefahrenhinweiskarten werden eher die Prozesse moduliert, Steinschläge, Murgänge; wo könnten sie in der Grössenordnung sein.
JB	*Und wie werden zukünftige Murgänge vorausgesagt? Werden ehemalige Ereignisse dazu verwendet?* *How are future debris flows predicted? Are past event extents used for predictions?*
E3	Ja, also bei der Gefahrenkarte ist es eigentlich so, sie blickt in die Zukunft aber nicht auf die nächsten 100 Jahren sondern die nächsten paar Jahre. Manchmal werden Gefahrenkarten auch wieder revidiert, wenn man merkt, es ändert sich etwas. Also es ist vorwärts, es ist auch… Gefahrenkarten sind auch ein Teil der Risikoanalyse, welche immer vorwärts blickend ist. In dem Sinne ist dies mit dem abgedeckt. Es gibt sicher auch Studien, bei welchen man das noch überlagert mit neuen Bauten etc. aber das ist nicht Teil der Schweizerischen Norm für die Gefahrenbeurteilung.
JB	*Welche Schwierigkeiten bestehen bei der Voraussage von Murgängen spezifisch?* *What are the difficulties when predicting debris flows / when generating hazard maps or planning measures of mitigation?*
E3	Ein schwieriger Punkt ist es, die Szenarien festzulegen. Man muss 3 Szenarien festlegen: ein 30-jährliches, welches alle 30 Jahre auftritt, ein 50-jährliches und ein 100-jährliches. Und diese Inputs festzulegen, das ist, glaube ich, etwas vom schwierigsten. Man muss das ganze Catchment gut kennen etc. und die Niederschlagsverhältnisse, oder, was es auch immer ist. Dort würde ich sagen, sind die grössten Unsicherheiten. Wirklich nachher im Prozess kann man das mit RAMMS ziemlich gut machen. Plus noch, im Feld muss man sicher gut schauen. Aber das, diese Szenarienunsicherheit mit den Inputwerten propagiert sich ins Modell, jetzt zum Beispiel oder in die Abschätzung im Perimeter der kartiert wird als Murgang-betroffen. Und ja, dann ist es natürlich sehr schwierig zu sagen, ob ein Haus betroffen ist oder nicht. Und weil es in der Gefahrenkarten dann scharf sein muss, ist man dann auch gezwungen. Es ist einfach in Planungsprozess in der Schweiz. Wir sagen immer, es wäre besser wenn man eine Unsicherheit reinbringen könnte, aber das ist nicht so einfach im Planungsprozess der Gemeinden, Kantone etc.
JB	*In der Unsicherheitsvisualisierung unterscheidet man zwischen einer zeitlichen Unsicherheit, einer Attribut-Unsicherheit und einer positionellen Unsicherheit. In welchem Bereich sind bei Murgängen die grössten Unsicherheiten anzutreffen?* *What are the main uncertainties about debris flow events? (positional vs. attribute uncertainty)*
E3	Also die zeitliche Unsicherheit gibt in dem Sinne nicht unbedingt, weil Gefahrenkarten nie voraussagen zum Beispiel: dieser Murgang ist in 3 Wochen, am Mittwoch in 3 Wochen. Dafür sind dann eher solche Weather-Forecasts, welche sagen: Schlecht Wetter, es könnte Murgänge geben. Bei der Attribut-Unsicherheit geht es in die Szenarien, dass man Inputs-Ding… das ist sicher wichtig. Und positionelle Unsicherheit ist dann, was ist betroffen oder?
JB	*Ja, also quasi eine räumliche Unsicherheit.*
E3	Das hängt einfach extrem zusammen. Das eine bedingt das andere.
JB	*In deiner eigenen Forschung, an welchen Aspekten dieser Unsicherheiten bist du am meisten interessiert? Oder welche denkst du, ist am wichtigsten zu erforschen?* *What aspects of uncertainty are you most interested in? (spatial, temporal attribute)*

E3	Von den Szenarien her, also eben, diese beiden sind fast nicht trennbar, die Szenarien und die Position, die positionelle (Unsicherheit). Also diese beiden brauche ich am meisten… dort, denke ich, würde es am meisten Sinn machen mit Unsicherheitsdarstellungen zu kommen.
JB	*Wo entstehen die grössten räumlichen Unsicherheiten?* *When are these uncertainties introduced?*
E3	Also wie meinst du, wo? Also im Kanton Bern oder wie?
JB	*Also nicht räumlich wo, sondern wo im Prozess von der Modellierung von zukünftigen Murgängen.*
E3	(…) Ich würde auch sagen bei der Definition der Szenarien. In dieser Range, der passieren kann. Und der Wahrscheinlichkeits… die eintreffende Wahrscheinlichkeitszuordungen. Dass man sagt mit diesen Grössen, mit diesen Input Szenarien, dass ist eben beispielsweise die 30-jährige Wahrscheinlichkeit, also die 30-jährige Return-Periode.
JB	*Du hast eben gesagt, dass in Gefahrenkarten keine Unsicherheiten dargestellt werden. Wenn das der Fall sein müsste, wie würdest du vorschlagen, die Unsicherheit von Murgängen in Karten darzustellen?* *How would you depict the positional uncertainty of debris flows in maps?*
E3	Also etwas muss ich noch präzisieren: Es gibt sozusagen die Residual-Hazard, dort werden Sachen, welche zum Teil weniger wahrscheinlich sind als 300 jährlich werden dort drin… Also wenn man das Gefühl hat, hier gibt es noch ein anderes Ereignis, welches passieren könnte, aber mit sehr tiefer Wahrscheinlichkeit, dann wird das weiss schraffiert dargestellt. Diese Möglichkeit gibt es. Und sag nochmal schnell wo…
JB	*Wie würde man die Unsicherheit visuell darstellen bei Gefahrenkarten oder auch spezifisch bei Murgängen?*
E3	Ja, ich würde, also bei… also was wir auch schon gemacht haben ist, zum Beispiel, dass man ganz viele… dass man die Inputparameter variiert. Und dann ganz viele Modellierungen laufen lässt und dann sieht, wo gibt es, ja… Wenn man zum Beispiel 100 Modellierungen laufen lässt, dort darstellt: eine Zelle oder ein Quadratmeter, durch wie viele dieser 100 Modellierungen ist sie betroffen. Entsprechend wenn es ganz viel ist, ist die Wahrscheinlichkeit hoch oder sonst weniger. So könnte… solche Sachen haben wir auch schon gemacht.
JB	*Wenn Visualisierungen erstellt werden, werden diese mehr unter Experten in der Wissenschaft gemacht oder werden sie auch mit Nicht-Experten geteilt?* *And are those debris flow visualizations used among experts and for experts of also for the public?*
E3	Dass man es mit Behörden teilt, kann sein. Aber das ist eher etwas innerhalb der Forschung, einfach im Planungsprozess. Es gibt vielleicht solche Pilotstudien aber im Planungsprozess ist es noch nicht möglich.
JB	*Hast du persönlich schon mal Schwierigkeiten gehabt über Unsicherheiten zu sprechen oder Unsicherheiten zu kommunizieren im Zusammenhang mit Naturgefahren gegenüber die Öffentlichkeit oder die Behören?* *And have you personally ever experienced difficulties when communicating uncertainties to decision makers?*
E3	Ja, es ist fast immer ein wenig… es ist fast immer ein Thema. Manchmal sagt man es auch transparenter oder nicht, also… wir machen in diesem Sinn keine Gefahrenkarten, das machen Umweltbüros und so. Wir sind eher in der Rolle der Experten und bringen uns ein. Ja, also es ist immer wieder ein Thema, die Unsicherheit von gewissen Prozessen und wie man damit umgeht. Wir arbeiten auch relativ viel im Ausland und dort gibt es teilweise andere Regeln bezüglich Gefahrenkarten. Aber es ist immer wieder ein Thema.
JB	*Denkst du die Darstellung von Unsicherheit in Murgang Gefahrenkarten wäre von Vorteil für Entscheidungsträger und Experten?* *Do you expect the depiction of uncertainties in debris flow risk maps to help decision makers and experts?*
E3	Ja, ich glaube schon. Also, man müsste dort sicher noch Arbeit leisten, dass sie damit umgehen können. Ich habe das Gefühl, langfristig läuft es eher in diese Richtung. Aber eben, gewisse Behörden sind dann nicht so

	ganz… also schauen es schon an, sind interessiert, aber schlussendlich für den Planungsprozess müssen sie zum Normalen zurück.
JB	*Hurricane cone und Konzept presentieren/present maps of hurricane cone and my sketch* *Meine erste Inspiration für meine Arbeit kam von der Hurricane Cone of Uncerainty. Diese sieht folgendermassen aus. Ich weiss nicht ob du diese schon gesehen hast. Es geht darum, dass diese dynamisch erstellt werden, wenn sich ein Hurricane ankündet. Dass man einen gewissen Zeitpunkt hat, welcher das 'jetzt' darstellt und man dann den höchstwahrscheinlichsten Verlauf darstellt. Und mit der Cone darum zeigt man eigentlich mögliche Bewegungsrichtungen.*
E3	Diese ist aber einfach weiss, diese ist nicht noch farbig zum Beispiel aussen wenig wahrscheinlich und so?
JB	*Nein, sie ist weiss, sie ist einfach pur und zeigt die 2/3 Wahrscheinlichkeit an basierend auf Erfahrungswerten von vergangenen Hurricanes. Das war meine Inspiration.* *Nun eine ganz schematische Skizze wie ich es mir vorstelle: Dass man an bekannten Murgang Standorten ähnlich darstellt, indem man auch sagt, wie du gesagt hast, man lässt viele Simulationen laufen und probiert die Wahrscheinlichkeit, dass ein Gewisser Ort betroffen ist, zu generieren. In diesem Beispiel arbeitet man mit der Transparenz: sprich Orte nahe am Fluss sind nicht transparent, sprich es sieht sicherer aus und je weiter man an den Rand kommt, desto transparenter wird die kartografische Darstellung.* *Denkst du das Konzept vom Hurricane Cone kann auf die räumliche Unsicherheit von Murgängen übertragen werden?* *Do you think the concept of the hurricane cone can be applied to the spatial uncertainty of debris flows?*
E3	Also ich verstehe nicht ganz, hier würde er los gehen (*zeigt auf Karte*) und dann… also was bedeutet das, es hat eine Farbabstufung von tiefen, also hier denkst du geht er los, er könnte auch weiter oben los gehen. Je blauer es wird, desto unsicherer ist es, dass er hier hin kommt?
JB	*Genau.*
E3	Und was ist die Abgrenzung?
JB	*Die Abgrenzung wäre auch entweder, wahrscheinlich ähnlich wie beim Hurricane Cone eine 2/3 Wahrscheinlichkeit.*
E3	Da kommt mir in den Sinn, ich habe einmal ein GIS Modell entwickelt, dort wurde auch mit qualitativer Wahrscheinlichkeit… ich kann dir das Paper schicken. Da wird mit qualitativer Wahrscheinlichkeit gearbeitet mit Farben. Also etwas in diese Richtung, jetzt das hier weiss ich nicht
JB	*Ja, das ist jetzt sehr schematisch.*
E3	Das glaube ich, ist nicht wahnsinnig nützlich, weil alles die gleiche Farbe ist. Aber wenn du mit anderen Farbkombinationen arbeitest, wo es wirklich anders ist, glaube ich schon. Also normalerweise, wie bei den Hurricanes, lassen sie das Modell ganz oft laufen und dann kann man solche Unsicherheiten auch aus… statistisch ausrechnen. Und das, also du müsstest mal suchen, ich kann dir da gerade nicht… wir haben es auch schon gemacht, ich schicke dir dieses Paper. Das Modell läuft nicht über ganz viele Modellläufe sondern es ist eine GIS Prozedur, welche ein wenig komplexer wäre. Aber grundsätzlich denke ich sicher, dass es nützlich ist.
JB	*Was sind typische Entscheidungen, welche mithilfe von Murgang Karten getroffen werden von Entscheidungsträgern?* *What are typical decisions, which are made with the help of debris flow event depictions?*
E3	Man braucht es manchmal, um zu entscheiden ob man irgendwo noch etwas bauen kann oder nicht. Oder was muss ich machen, wo muss ich eine Massnahme treffen, einen Damm bauen oder ein Rückhaltebecken. Das sind so typische Fragen wie mit den Gefahrenkarten.
JB	*Noch zum letzten Thema, welches ich besprechen möchte, da geht es um den Einfluss des Klimawandels:*

	Welche Veränderungen erwartet man bei Murgängen in der Schweiz aufgrund des Klimawandels? *How are the characteristics of debris flows in Switzerland expected to evolve due to climate change?*
	At this point of the interview the recording stopped due to the overheating of the recording device. The last three questions were answered by email by the interviewee after the interview.
E3	Bis anhin konnten noch keine markanten Veränderungen bei Murgängen festgestellt werden. Für die Zukunft deutet vieles auf höhere Frequenz und/oder Intensität (also Grösse) hin, und zwar weil klimatische Extremereignisse (Niederschlag, Temperatur) häufiger und wohl intensiver werden, weil im Hochgebirge mehr Schutt verfügbar wird durch Rückzug der Gletscher und Auftauen des Permafrosts.
JB	*Welche Unsicherheiten bestehen bezüglich der Vorhersage zukünftiger Murgänge?* *What are the uncertainties tied to future debris flow predictions?*
E3	Die Frage, wann ein Murgang geschieht, ist immer unsicher. Zudem können auch an neuen Orten Murgängen entstehen, insbesondere im Hochgebirge wegen den starken Veränderungen dort (Kryosphäre).
JB	*Gibt es etwas, was wir bisher nicht erwähnt haben, worüber du noch sprechen möchtest?* *Is there any aspect of debris flows and their uncertainties that we have not touched on and that you would like to mention?* *Input given in email: Dein Input bezüglich der Machbarkeit die Unsicherheiten kartografisch zu Integrieren im Prozess der Gefahrenkartenerstellung, aber mittelfristig wird nicht erwartet, dass Unsicherheiten Entscheidungsträgern gezeigt werden*
E3	Genau, von behördlicher Seite her brauchen sie weiterhin die genauen Abgrenzungen der Gefahrenzonen zur Planung. Aber Karten, die Unsicherheiten darstellen, könnte ich mir gut vorstellen als zusätzliche Info, die in den Prozess einfliessen kann.

D. Transcript of Expert Interview E4

Interviewer:	**Interviewee:**	**Role:**	**Date:**
Jana Bracher	E4	Practitioner	21.09.2021

This interview was conducted in Swiss German and translated to German during the transcription which resulted in the rephrasing of certain statements. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability.

JB	*Damit es auch protokolliert ist: Ist es für dich in Ordnung, wenn ich das Interview aufnehme?* *For the protocol: Is it okay for you if I record the interview?*
E4	Ja.
JB	*Perfekt, dann fangen wir an. Erkläre bitte kurz, was dein daily-business in der Arbeit mit Naturgefahren ist.* *Then, let's begin with the questions. Please explain some of the everyday tasks of your position regarding natural hazards.*

E4	(*Beschreibung Arbeitsort & Position*) Wir haben die Aufgabe, die Gemeinden in ihren wasserbaulichen Aufgaben zu beraten, zu unterstützen. Wir haben aber auch die Oberaufsicht. Das heisst, wenn etwas nicht so läuft, oder so gemacht wird, wie es gemacht werden sollte, dann müssen wir auch eingreifen. Wir betreuen den Gewässerunterhalt der Gemeinden. Wir subventionieren die Wasserbaugeschäfte, welche sie verwirklichen, sehr hoch. Wir leiten die Bundesgelder weiter an die Gemeinden und verwalten sie auch in dem Sinne. Wir bürgen für die Bundesgelder. Und zuletzt haben wir noch die Wasserbaupolizei-Funktion und wir beurteilen Bauvorhaben in Gewässernähe oder in Überflutungsgebieten. Was die Abstände anbelangt, ermöglichen wir, dass die Gewässer zugänglich sind, dass nicht zu nahe an Gewässern gebaut wird. Oder, dass auch so gebaut wird, dass im Falle von Überflutungen nicht zu grosse Schäden entstehen. Und dazu gehören auch die Gefahrenkarten, welche ein Instrument sind, aber auch periodisch aktualisiert werden. Das ist auch noch ein grosser Aufgabenbereich.
JB	*Das leitet mich gleich zur nächsten Frage. In deiner Arbeit: Was ist der Zweck der Gefahrenkarten? This leads to the next question. In your work: What is the purpose of the hazard map?*
E4	Eine Gefahrenkarte soll bezwecken, dass man sich der Gefahr bewusst ist, dass man überhaupt weiss, wo ist welche Gefahr vorhanden. Mit der einfachen Abstufung von vier Kategorien: Restgefahr, kleine, mittlere und hohe Gefahr. Das ist schon der erste wichtige Schritt: das Bewusstsein der Gefahr. Das war lange Zeit nicht der Fall. Für uns als Geographen ist es oftmals offensichtlich. Draussen im Feld ist es nicht immer offensichtlich. Und wenn man überhaupt eine Gefahr erkannt hat, geht es darum, Rückhaltemassnahmen zu treffen. Sei es an bestehenden Objekten, dass man dort Massnahmen trifft, um direkt am Objekt zu schützen oder an den Gerinnen selbst, wenn es darum geht, viele Objekte vor einer bestimmten Gefahrenquelle zu schützen. Wenn man an neuen Orten arbeitet in Bereichen, welche überflutungsgefährdet oder naturgefährdet sind - bei uns sind es natürlich die Überflutungs- und die Murganggefahr. Wenn jemand neu bauen oder umbauen möchte in einem solchen Gefahrenbereich, geht es darum, dass wir die Auflagen so definieren, dass die Gebäude entsprechend robust sind und primär keine Menschenleben, aber auch keine höheren Sachwerte gefährdet sind.
JB	*Du hast es bereit angesprochen, vielleicht noch explizit: Welche Entscheidungen können mithilfe der Gefahrenkarten getroffen werden?* *You already mentioned it, but maybe the explicit question: What decisions can be made with the help of hazard maps?*
E4	Wenn es um die konkreten Vorhaben geht, macht man im Prinzip zuerst einen Verschnitt zwischen Objekt und Gefährdung. Das heisst, man muss sehen, was man macht. Der nächste Schritt ist meist… Die Kategorien der Gefahrenkarten sind sehr grob. Man hat die Intensitäten gering-mittel-hoch. Gering bedeutet bis 0.5m (Ablagerung), mittel 0.5 bis 2m und hoch ist mehr als ein Meter. Und dann noch die Geschwindigkeitshöhe, wobei v*h relevant ist. Aber allein aufgrund der Gefahrenkarte und der Intensitätskarte wissen wir noch nicht viel über die effektive Gefährdung. In vielen Fällen heisst es, die Gefährdung ist mittel - also blau - das ist oft eine mittlere Intensität. Es kann auch eine geringe Intensität mit einer hohen Wahrscheinlichkeit sein, die zu blau führt. Aber in sehr vielen Fällen weiss ich im blauen Bereich folgendes: wenn das Wasser nicht zu schnell fliesst, ist es zwischen 0.5 und 2m hoch. Und für den Bauherrn ist das natürlich eine riesige Spanne. Und er muss eigentlich genau wissen, bis wo er sein Objekt schützen muss. Und dort muss… Entweder haben wir an gewissen Orten bereits bessere Daten oder Gutachten im Umkreis, wo wir eine genauere Aussage machen können, und dann teilen wir das mit. Und aufgrund dieser Angaben muss der Planer, also der Architekt oder Ingenieur, Objektschutzmassnahmen treffen. Und wenn wir es nicht genau sagen können, dann muss der Bauherr ein Objektschutzgutachten machen. Der erste Schritt ist eigentlich das Feststellen der genauen Hochwasserschutzquote. Dass man in Metern über Meer - am besten oder

	gegenüber dem Terrain – weiss… Ich sage jetzt mal, die Schutzmassnahmen über 75cm, respektive ca. so viele Meter über Meer, muss man sicherstellen. Es kann dann auch noch je nach Fassade variieren. Das ist nicht immer gleich.
JB	*Wie gestaltet sich dein Entscheidungsprozess? Welche Akteure sind involviert? Welchen Stellenwert hat die Gefahrenkarte in deiner Entscheidung?* *How does your decision-making process take place? Which actors are involved? And how high is the relevance of the hazard map?*
E4	Die Gefahrenkarte ist eine wichtige Grundlage, schon nur dass es (ein Fall) zu uns kommt. Die Baugesuche kommen über die Gemeinde und bei uns gibt es noch die Regierungsstadtämter. Sie schauen das zuerst an und müssen überhaupt erst feststellen, dass es ein Problem aufgrund von Naturgefahren gibt. Und wenn sie das nicht feststellen können, kommt es gar nicht zu uns. Für sie ist die Gefahrenkarten das Instrument. Also sie schauen anhand der Gefahrenkarte, ob das Vorhaben eine rote oder blaue Gefahrenzone betrifft. Und wenn sie etwas Erfahrung haben und es gut machen, müssten sie auch in der gelben Zone schauen, ob es ein sensibles Objekt ist, welches in der gelben Zone liegt. Also ist es eine Einstellhalle, ein Altersheim oder ein Schulhaus. Dann müssen sie es uns auch schicken, wenn es in der gelben Zone ist. Das ist eigentlich der erste Schritt und wenn es einmal bei uns gelandet ist, definieren wir alle weiteren Schritte mittels Auflagen, welche wir letztendlich… Oft ist es so, dass wir ihnen sagen, was sie machen müssen. Aber eigentlich müsste die Bauherrschaft die Lösung präsentieren. Und das funktioniert nicht immer ganz so gut, es kommt ein wenig auf den Gutachter an, falls sie einen haben. Oft geht es auch noch darum, dass man diskutiert, ob man Dammbalken in den Türen einbauen kann als mobile Massnahme, wenn man einen Meter über dem EG das Schutzniveau definiert. Das sind die Punkte, welche immer wieder zu diskutieren sind. Das machen wir abhängig von der Prozessquelle, der Reaktionszeit und aufgrund von der Erfahrung. Und die Machbarkeit der Gebäude, was ist überhaupt machbar.
JB	*Fördert die Gefahrenkarte in der heutigen Form deine Entscheidungen oder gibt es auch Herausforderungen? Du hast erwähnt, dass die Abstufung in den Gefahrenkarte ziemlich grob ist.* *Does the hazard map in its form today support your decision-making or are there also challenges? You mentioned, the broad classification in the hazard map.*
E4	Seit einer Weile machen wir neben den Gefahrenkarten noch Zusatzprodukte, wie es andere Kantone bereits machen. Beziehungsweise wir lassen sie machen. Ich muss auch sagen, die Gefahrenkarten machen eigentlich offiziell die Gemeinden und wir unterstützen sie und subventionieren schlussendlich 90%, sprich die Gemeinden zahlen nur noch 10%. Aber sie sind eigentlich der Auftraggeber. Aber vom fachlichen sind wir natürlich extrem eng involviert und die Gemeinde macht keinen Entscheid, ohne, dass wir gesagt haben, was sie machen müssen. Letztendlich sind wir dort fast federführend, nicht offiziell, aber faktisch schon. Und wir empfehlen immer, die sogenannten Fliesstiefenkarten wie auch die Fliessgeschwindigkeitskarten zu machen. Und aus diesen kann man dann auch Hochwasserschutzquotenkarten ableiten, das ist sozusagen das Produkt aus den beiden. Es gibt die Energiehöhe, das sind Karten mit Höhenlinien von den… Es gibt verschiedene Produkte. Man kann den Wasserspiegel nehmen oder effektiv die Schutzquote, der Unterschied ist dabei den Freibord, respektive die Energiehöhe. Wenn das Wasser schnell fliesst, dann haben wir eine höhere Schutzquote. Demnach ist die Schutzquote nicht auf der Wasseroberfläche, sondern höher, weil das Wasser noch aufbranden kann. Dort fügt man den Freibord hinzu.
JB	*Gibt es auch Mittel oder Parameter spezifisch für Murgänge? Weil das eher spezifisch für Hochwasser war.* *Are there also tools and parameters specifically for debris flows? Because that was more specific to floods.*
E4	Ja, wobei Murgänge sind vom Prozess her das gleiche. Es ist einfach eine andere Rheologie. Er ist dickflüssiger, hat also eine andere Dichte und die Viskosität spielt noch eine Rolle. Aber letztendlich, also es

	ist eben... Bei uns im Gebiet sind Murgänge nicht so relevant. Das sind eher noch die Hangmuren, das sind sehr flachgründige Rutschungen. Aber da sind wir nicht zuständig. Das ist das Spezielle bei der Behörde, es ist immer klar definiert, welche Abteilung, wofür zuständig ist. Und bei den Hangmuren ist es die Abteilung Naturgefahren, welche zuständig ist. Aber auch die Hangmuren kann man modellieren. Es ist ähnlich wie beim Wasser. Ein Hochwasser ist ein flächiger Prozess und hat natürlich viel mehr betroffene Gebiete oder Objekte. Das ist deutlich die Prozessquelle, bei der am meisten Schadenpotential existiert. Bei der Mortalität ist es anders, da sind es die Lawinen. Sprich bei den Schneelawinen haben wir am meisten Tote pro Jahr. Bei den Hochwassern haben wir zum Glück nicht so viele. Die Schäden sind dafür umso höher.
JB	*Du hast kurz angesprochen, dass die Gefahrenkarten von den Gemeinden erstellt, beschreibe bitte trotzdem schnell das Vorgehen, wenn Gefahrenkarten professionell erstellt werden?* *You mentioned that the municipalities have to create the hazard maps. Could you still briefly explain the process of creating the hazard maps?*
E4	Es sind noch Schritte, welche rein administrativ sind. Also die Auswahl eines Büros, dass man verschiedene Offerten einholt und sie auswertet. Das sind klar definierte Prozesse. Vorher muss man noch ein Pflichtenheft machen, das ist ein ca. 60-80-seitiges Dokument. Schon das zeigt, allein das Pflichtenheft zu erstellen, was eigentlich auf Seiten der Gemeinden geschieht, ist eine ziemliche Sache. Es ist unterdessen so aufwendig, dass wir in der Regel dort schon ein Büro beiziehen, welches sie unterstützt, weil wir die Kapazität auch nicht mehr haben, um das zu machen. Dort geht es darum zu definieren, was man wo genau macht. Und aufgrund dieses Pflichtenheftes machen die Büros die Offerten und da wird klar definiert, welche Prozesse betrachtet werden, wie genau man es betrachtet, ob man die Hydrologien überarbeitet, wie ist die Hydraulik und so weiter. Ist es eine Teilrevision oder eine Gesamtrevision? Welche Prozesse werden wie tief behandelt? (…) Wenn die Startsitzung stattgefunden hat, dann beginnt das Büro. Im Bereich Hochwasser wird auch im Pflichtenheft definiert, was ist gutachterlich gemacht und was modelliert wird. Und gutachterlich heisst dann wirklich schon aufgrund von Karten, vielleicht werden im Feld noch Kartierungen gemacht und nach Erfahrung diese Prozessgebiete kartiert. Das haben wir in Lauterbrunnen auch gemacht. Und das andere ist das Modellieren, wobei man es wirklich rechnet. Und man muss aufpassen, es ist nicht immer besser. Beide Methoden haben ihre Vor- und Nachteile. Aber wenn man es flächig betrachten will, wenn man grössere Austritte vom Gerinne in die Ebene hat, dann geht es inzwischen nicht mehr, ohne zu modellieren. Das erreicht man gutachterlich nicht, keine Chance.
JB	*Was sind die grössten Herausforderungen beim Erstellen von Gefahrenkarten?* *What are the biggest challenges when creating hazard maps?*
E4	Was man immer wieder diskutiert oder schwierig ist, ist einerseits die Hydrologie oder die Abflusswerte, welche verwendet werden. Es gibt sehr viele Methode, um diese festzulegen. Und es ist schwierig, weil wir machen das meist pro Gemeinde, aber auch innerhalb einer Gemeinde gibt es Teilrevisionen. Der Umgang mit Veränderungen ist dann noch schwierig. Wenn man feststellt… Unsere Gefahrenkarten sind in der Regel anfangs der 2000er Jahre erstellt wurden, die sind jetzt etwa 20-jährig. Wir revidieren pro Jahr etwa 5-6 Karten. Da sind zum Teil andere Ansätze gewählt worden. Man hat heute schon tendenziell höhere Abflusswerte. Das führt oft zu Diskussionen, wie man damit umgeht, sodass das Ganze in sich zusammenpasst. Auch benachbarte Gemeinden oder man hat zum Teil auch Gemeindegrenzen, welche nicht den Einzugsgebieten entsprechen. Sie (die Gefahrenkarten) sind manchmal ein wenig an ihnen (den Einzugsgebieten) orientiert. Der Umgang mit den Schnittstellen ist noch schwer zu definieren. Man hat immer irgendwo irgendwelche Lücken, die man nicht vermeiden kann. Zumindest für ein paar Jahre bis zur Revision der unter- oder obenliegenden Gemeinde, dann kann man es aufeinander abstimmen. Das ist noch ein grosser Aufwand, welcher über uns läuft, weil niemand ausser uns den Überblick über die Kommune

	hinaus hat. Natürlich müssen auch die Büros links und rechts schauen, aber das ist schon schwierig. Wir hatten auch ein Projekt, bei dem man für einen Bach eine Hydrologie festgelegt hat und zwei Jahre später machte man die Gesamtrevision und diese Gutachter oder dieses Büro wählte eine andere Methode. Man kann nicht sagen, dass eine Methode richtig oder falsch ist, es ist einfach eine andere Methode und das führt zu leicht anderen Resultaten. Und dann passt ein einzelner Bach nicht mehr ins Gesamtbild. Da muss man einfach wissen, wie man damit umgeht. Man baucht eine gewisse Planbeständigkeit. Man kann nicht 1 bis 2 Jahre nachdem man irgendwo ein Hochwasserprojekt gemacht hat, wieder kommen und sagen, jetzt reicht es nicht mehr aus, auch wenn es ein kleiner Bach ist und die Gefahr dort zu einem gewissen Grad bis zum 100-jährlichen Ereignis behoben ist. Und da machen wir uns als Behörde dann total… Da werden wir nicht mehr verstanden. Und dort müsste man sicherstellen, dass man die Objektivität und die Realität auch die Nachvollziehbarkeit, die Planbeständigkeit, alles unter einen Hut bringt. Das ist nicht immer alles konsistent. Da müssen wir Kompromisse eingehen.
JB	*Das mit den unterschiedlichen Methoden ist sicher eine Quelle von gewisser Unsicherheit beim Erstellen der Gefahrenkarten. Gibt es noch andere Unsicherheiten vielleicht auch noch spezifischer bei Murgängen?* *The different methods are certainly a source of uncertainty. Are there other uncertainties? Maybe also specific to debris flows?*
E4	Es gibt enorme Unsicherheiten. Wir sagen auch immer wieder, wir haben gerade einen Fall, bei dem man diskutiert… Bei einem Projekt, bei dem die Geometrien klar definiert sind, haben wir festgestellt, dass man an einem Ort die Geometrie nicht erreichen kann. Es ist ein Brückendurchlass, bei dem alles einbetoniert ist, den kann man nicht mit verhältnismässigem Aufwand so weit öffnen, dass so viel Wasser durchkommt, wie man möchte. Und in diesem Fall setzt sich der Ingenieur jetzt auf den Standpunkt: Hier können wir nicht die Wassermengen abführen, welche wir gerne hätten. Und das ist absolut, um das 100-jährliche Schutzziel zu erreichen. Aber in diesem Moment muss man schnell reagieren, das ist eine laufende Baustelle, man kann nicht die Hände verwerfen und sagen: Was machen wir jetzt? Das können wir nicht erreichen. Letztendlich ist es immer noch… In diesem Fall haben wir sehr grosse Verbesserungen gegenüber heute und da muss ich auch sagen, dann erreicht man eben nur 90% vom Abfluss, den man möchte, also nicht ganz das 100-jährliche. Das 100-jährliche ist an einen Wert X gebunden, beispielsweise 20 Kubikmeter pro Sekunde. Wenn da nur 18 durchkommen, ist es natürlich schade um die 2, welche weniger sind, aber wenn heute nur 10 durchkommen und neu 18, dann ist es trotzdem fast eine Verdoppelung. Da sage ich, das ist eine wesentliche Verbesserung. Und ich plädiere immer dafür: Wir brauchen die Jährlichkeiten 30-, 100-, 300-jährlich auch für das Politische zur Kommunikation. Aber das war schon in meiner Dissertation (*Beschreibung Dissertation*). Das sind dann Szenarien. In der Hydrologie haben wir das Glück, dass man bei gewissen Bächen Messreihen hat, welche 100 Jahre abdecken zum Teil etwas mehr, zum Teil etwas weniger. Oder mehrere Jahrzehnte. Da kann man eine Statistik machen, man kann die Jährlichkeiten einigermassen ausrechnen. Aber auch dort ist die Zeitspanne, auf welche man zurückblickt, sehr kurz. Das würde voraussetzen, dass die Zukunft gleich ist wie die Vergangenheit. «The past is the key to the future.» Die Vulkanologen, das ist wichtig in den Geowissenschaften. Das ist zu einem gewissen Grad so, aber wenn du… Wir sind nicht in einem statischen System. Wir sind in einem dynamischen und es gibt Veränderungen. In solchen Fällen plädiere ich darum dafür eher etwas… Man kann auch mal von Klein-, Mittel- und Grossereignissen sprechen anstatt immer von 30-, 100- und 300-jährlich. Und vor allem muss man auch die Verbesserungen sehen und nicht nur das, was man nicht erreicht. Und abgesehen davon, auch in diesem Fall hier, kam der Ingenieur und meinte, man habe im Vorprojekt zu hohe Strickler-Werte verwendet. Das bedeutet eine zu feine Rauigkeit, sprich es würde eine zu hohe Geschwindigkeit suggerieren als tatsächlich präsent ist. Das heisst also, die Kapazität wäre eigentlich noch kleiner als erhofft und das ist absolut richtig, aber dort… Die Ingenieure wollen einfach etwas mit fixen Zahlen berechnen und dort ist zu wenig Gefühl für

	Unsicherheit vorhanden. Dann müssen wir sie manchmal auch auf den Boden der Tatsache holen und sagen: Lasst die fünf geradestehen. Allein wenn ich den Querschnitt betrachte und wie er vergrössert wird, muss ich sagen, es ist eine gute Sache.
JB	*Und wird Qualität der verwendeten Daten im Prozess der Gefahrenkartenerstellung berücksichtigt?* *Is the quality of the data considered when creating hazard maps?*
E4	Nein. Nein, schlicht nein. Keine Zeit. Für solche Sachen hat kein Büro Zeit. Sie sind derartig unter Druck, das schnell zu machen. Sich über Unsicherheiten Überlegungen zu machen, bleibt auf der Strecke. Es ist aber auch nicht so schlimm. Wichtig ist, dass wir uns dessen im Vollzug bewusst sind. Und auch bei den Gefahrenkarten: Wir haben im Verlauf des Erarbeitens der Gefahrenkarten verschiedene Sitzungen und da kommen immer wieder solche Fragen auf. Und dort können wir Gegensteuer geben, wenn es in eine Richtung ginge, welche nicht gut wäre.
JB	*Sprich in der Gefahrenkarte werden die Unsicherheiten oder die räumliche Unsicherheit der Phänomene nicht explizit thematisiert?* *So, uncertainties or spatial uncertainties are not explicitly treated in the hazard maps?*
E4	Nein null, dürfte sie auch nicht. Wir haben die Gefahrenmatrix, die 3x3 Matrix mit den neun Feldern, welche… Diese sind schwarz-weiss. Das ist das System und wir betrachten im Hintergrund die Fliesstiefenkarten und die Abflussgeschwindigkeitskarten. Diese sind weniger schwarz-weiss, aber sie suggerieren auch eine Genauigkeit, welche nicht vorhanden ist. Aber man sieht zumindest die Übergangsbereiche. Sonst weiss ich nicht, wie abrupt der Übergang zwischen gelb zu blau ist. Oder wenn man die Intensitätskarten mit diesen Stufen betrachtet. Auch dort muss man sich bewusst sein, denn ein Projekt muss eine Kostenwirksamkeit nachweisen. Und die Kostenwirksamkeit heisst nichts anderes, als dass aus jedem investierten Franken mindestens ein Franken Nutzen gezogen werden muss. Dafür gibt es ein Tool, welches der Bund vor 15 Jahren zu entwickeln begonnen hat und jetzt verfeinert. Das sogenannte EconoMe. Vielleicht hast du davon schon gehört. Dort hat man auch klare Szenarien. Im Standardszenario hat man das 30-, 100- und 300-jährliche Ereignis und für das braucht man die Intensitätskarten und derjenige, der das ausfüllt, macht nichts anderes als die Intensitätskarte beispielsweise für das 100-jährliche zu betrachten. Das ist eine Fläche, welche zeigt, was alles betroffen ist. Und dann fängt er an, Häuschen zu zählen. Es sind beispielsweise 20 Einfamilienhäuser, 5 Zweifamilienhäuser und noch einige Mehrfamilienhäuser. Diesen sind Werte hinterlegt. Dann können sie auch noch Strassen, also Infrastruktur, oder Sonderobjekte definieren. Und das sind dann alles Werte, welche einfliessen. Und dann weiss das System, wie viel Schadenpotenzial vorhanden ist. Das wird verschnitten mit der Intensität und der Wahrscheinlichkeit, wie ihr es gemacht habt. Das gibt dann das jährliche Schadenpotenzial, woraus man dann das Risiko ausrechnen kann. Wenn man dort unscharfe Karten hätte, könnte man das gar nicht mehr machen. Also braucht es dort die Schärfe, auch wenn sie nicht ganz richtig ist.
JB	*Und für Entscheidungsträger/innen würdest du sagen, wenn die Unsicherheit dargestellt würde, würde es ihnen helfen oder würde es die Entscheidung erschweren?* *For the decision-makers: would you say depicting uncertainties would support or complicate their decisions?*
E4	Ich möchte nicht deine Arbeit in Frage stellen, ich finde es wichtig, dass man… Unsicherheit ist ein riesiges Thema. Aber in der Praxis empfehle ich nicht, dass man zusätzliche Unschärfen darstellt. Ich sehe das eher als wissenschaftlichen Aspekt, der auf einer gewissen Stufe berücksichtigen muss, aber wir haben jetzt schon… Mit unseren Gefahrenkarten kommen die Gemeinden einigermassen zurecht. Aber sie sind bereits mit den Zusatzprodukten überfordert. Und wenn man dort noch Abstufungen reinbringt, bringt das gar nichts. Also erschwert es für sie nur die Triage. Auf unserer Ebene ist es auch nicht, wie soll ich sagen. Es ist noch schwierig. Vielleicht ist es beim Murgang noch etwas anders. Wobei ich auch dort denke, dass es eine

	gewisse Einfachheit braucht, damit man es in der Praxis umsetzen kann. Diese ist im Moment weitgehend gegeben und das sollte man nicht zusätzlich komplizierter machen. Aber um die Prozesse zu verstehen, auch um das Risiko zu verstehen, ist es absolut richtig, dass man das auf der Ebene der Forschung sieht, was man machen kann. Und man kann dort vielleicht gewisse Tools entwickeln, aber ich denke der Nutzerkreis ist sehr klein.
JB	*Wenn man die Unsicherheiten integrieren würde, würde das den Prozess des Erstellens der Gefahrenkarten komplizierter machen?* *If one were to include uncertainties, would it make the process of producing a hazard map more complicated?*
E4	Ja, es ist natürlich die Frage, wie man es integriert. Meine Meinung ist, dass man nichts an der konzeptionellen Geschichte der Gefahrenkarte ändern darf. Das wäre verheerend. Ich denke dort… Das ist ein anderes Thema: Das BAFU tendiert dazu, immer noch etwas weiterzukommen, noch etwas zu verbessern. Aber wir sind an einem Punkt angelangt, an dem man derartig viele Informationen hat, welche nicht mehr… entscheidend ist nicht, ob eine Information etwas besser oder schlechter ist, sondern wie kommt sie überhaupt an? Werden die richtigen Schlüsse daraus gezogen? Die Gefährdungskarte des Oberflächenabflusses ist ein sehr gutes Instrument. Sie zeigt nicht die Unsicherheiten an, aber sie ist relativ fein, sie ist nicht stark generalisiert, denn die Gefahrenkarte ist stärker generalisiert. Und das ist für uns sehr gut zum Arbeiten, eine sehr gute Information. Aber man muss auch dort sehen: Die Planer sind damit überfordert. Es braucht Unterstützung von Leuten wie uns, um sie zu interpretieren. Und letztendlich ist es effektiv so, dass die Planer schon so viele Anforderungen haben, welche sie berücksichtigen müssen. Sie müssen eine Checklist haben, auf der sie schnell Sachen abhaken können. Und sich nicht noch mit… Sie können sich gar nicht noch mit Details beschäftigen. Sie müssen einfach konkret wissen, wenn es um Hochwasser oder einen Murgang geht, wenn sie… Bei mir ist das Ziel erreicht, wenn sie realisiert haben, dass sie ein Problem haben und sie darauf eingehen. Sie versuchen ihr Objekt so anzupassen, dass diese Gefährdung wesentlich reduziert oder am besten behoben wird. Aber dort muss man Abstriche machen, auch was die Unsicherheit anbelangt. Es ist sehr pragmatisch. Manchmal sage ich auch einfach aufgrund der Erfahrung im blauen Bereich, dass ich weiss, es ist 0.5 bis 2m Ablagerung. Ein Gutachten zu machen kostet 5'000 Franken. Je nach Bauherrschaft, weiss ich, dass das gar nicht… Es führt zu nichts, wenn ich da ein Gutachten verlange. Dann sage ich manchmal: seht zu, dass ihr wasserdichte Türen einbaut, welche möglichst einem halben Meter Wasserdruck standhalten und dann haben wir mehr erreicht, als wenn wir «hin- und her-gutachten». Es braucht da einen gewissen Pragmatismus. Ausser dann mit diesen Produkten… Du würdest stauen, wenn du ein Gefahrenkartendossier sehen würdest, das sind «Monster-Dossier», das ist sehr umfangreich. (…) Also es ist wirklich sehr aufwendig und es sind sehr viel Informationen. Das kann eine Gemeinde gar nicht mehr handeln, es ist schon viel.
JB	*Vielleicht noch etwas zur Standardisierung: Beispielsweise das Farbschema ist stark standardisiert in rot-blau-gelb. Von der kartographischen Perspektive ist es ein Farbschema, welches sehr arbiträr definiert ist und nicht kartographischen Standards entspricht. Wenn es nicht standardisiert wäre… (*Telefon läutet*, kurze Pause) Wegen des Farbschemas, wenn es nicht so standardisiert wäre, würdest du es ändern, oder würdest du es beibehalten?* *Regarding the standardisation: The colour scheme for example is highly standardized with red-blue-yellow. From a cartographic perspective this arbitrarily defined colour scheme does not conform with cartographic principles. If it was not standardized… (*phone rings, short break*) Regarding the colour scheme, if it were not standardized would you change it or keep it that way?*
E4	Ich sehe keinen Grund, es zu ändern. Die synoptische Gefahrenkarte ist schon zu wenig Information. Ich brauche viel mehr Information, das ist der Grund, weshalb wir die Fliesstiefen- und

	Fliessgeschwindigkeitskarten wollen und auch pushen. Aber für draussen, für die Triage, für die Behörde auf kommunaler Ebene ist es zwingend, dass sie so ein einfaches Tool haben. Das ist eine Vorstufe. Mir ist wichtig, dass man die Berechnungen - nicht das Gutachterliche, sondern das Modellieren, welches man speziell bei den Wasserprozessen immer mehr macht - dass man diese Resultate in einer guten Form aufbereitet und man diese auch noch in einigen Jahren oder in 10 Jahren nutzen kann. Modelliert wird bereits seit einer Weile, aber die Modellresultate hat man direkt in die Intensitätskarten und anschliessend in die Gefahrenkarten übersetzt und dadurch ist enorm viel Information verloren gegangen, welche nicht mehr zur Verfügung steht. Und das ist schade. Mir geht es darum, die Informationen, welche man inzwischen meist sowieso generiert, zu sichern. Das ist ein Modell-Lauf mit gewissen Annahmen, aber das ist aus meiner Sicht… Ich kann diese Unsicherheiten dort ein wenig reininterpretieren. Wobei man schon von dem, was man modelliert, ausgeht. Aber die Unsicherheiten decken wir situativ mit Freiborden ab. Bei Überlegungen wie: Ist hier Schwemmholz oder eine Auflandung / Übersarung ein Problem? Und je nach dem wählen wir den Freibord etwas höher oder tiefer.
JB	*Zwei Optionen für die Entscheidungsaufgabe: Studienteilnehmende sind Gemeindevertreter im Gefahrenmanagement.* *Welches Haus (A oder B) hat im Evakuationsplan eine höhere Priorität?* *Welches Haus (A oder B) sollte eher baulich geschützt werden?* *Hier wollte ich nach deiner Meinung fragen, welcher Aufgabentyp wäre realistischer im Kontext der Gefahrenkarte und welche Entscheidungen damit getroffen werden?* *There are two options for my decision task: study participants will be municipality representatives in hazard management* *Which of two houses should have a higher priority in the evacuation plan?* *Which of two houses should be protected with structural measures?* *I want to ask about your opinion, which of the two decision tasks would be more realistic?*
E4	Sie zielen auf zwei verschiedene Tools ab. Das eine geht in Richtung Notfallplanung, das sind die konzeptionellen oder organisatorischen Massnahmen, welche die Gemeinde, also die Wehrdienste treffen. Allenfalls Evakuation… Dort Karten mit einer Wahrscheinlichkeit der Betroffenheit umzusetzen ist schwierig. Also ob man sich das vorstellen kann… Die Notfallplanung ist auch eine ziemlich grosse Aufgabe. Auch dort erfolgt die Abgrenzung sehr schnell und pragmatisch aufgrund von Erfahrung, aber natürlich auch der Gefahrenkarten. Aber ich bezweifle, dass man dort Zeit hat, irgendwelche weitere Tools zu schaffen. Die Frage ist, ist es vorhanden? Ist es flächendeckend? Wo hat man Zugang dazu? Dann vielleicht noch eher… Aber ich denke schon, dass man die Gefährdungskarte Oberflächenabfluss auch in der Notfallplanung immer wieder pushen muss und sagt: schaut sie euch an, denn sie gibt auch noch Hinweise. Die geht auch immer wieder vergessen. Das ist auch Information, welche schwierig zu handeln ist.
JB	*Es geht mehr um mein Experiment, denn es gibt viele Experimente mit Entscheidungen, welche aus der Luft gegriffen sind und nicht realistisch oder realitätsgetreu sind. Das möchte ich verhindern und möchte eine Entscheidungsaufgabe stellen, welche auch realistisch ist.* *It's more about my experiment. There are a lot of experiments with unrealistic decision-making tasks. I want to prevent that and choose a realistic decision-making task.*
E4	Es ist schon nicht so, dass diese Triage, in der man allenfalls Massnahmen trifft, nicht realitätsnah ist, im Gegenteil, das ist… Durch das Einbinden der lokalen Leute wird weitgehend sichergestellt, dass man nahe an der Realität ist. Das Schwierige ist bei den Extremereignissen, bei welchen die Erfahrung fehlt. Dort existiert eine Lücke. Ich bin sehr vorsichtig, wenn es um Extremereignisse geht. Da müssen wir nicht viel machen in der Schweiz, weil das ist überall… Es gibt nicht viele Orte, welche nicht von Extremereignissen

	betroffen sein könnten. Damit müssen wir umgehen und da geht es auch um Risikoakzeptanz und Risikokultur. Aber noch zum anderen, was du erwähnt hast, zu den baulichen Massnahmen. Es ist auch noch schwierig… Es geht in die Richtung: Der Bund, also das BAFU, möchte Gesamtplanungen pushen. Das hat nichts anderes zum Zweck als zu untersuchen, wo die grösseren… Sprich die Risikokarten, welche sie möchten, welche wir bald machen müssen, zeigen die Hotspots. Und basierend darauf müssen wir die Gesamtplanungen machen. Das beinhaltet die Priorisierung, bei der man zuerst jene Punkte angeht, welche «Klumpenrisiken» sind. In der Theorie ist das nachvollziehbar. In der Praxis funktioniert es null und gar nicht. Weil du kannst nicht einfach… Also, doch du kannst mit den Steuern zu lenken probieren. Aber unter Umständen stösst man auf solch grosse Widerstände, dass es einfach nicht effizient ist, Mittel einzusetzen. Und da bin ich klar der Meinung, dass man bei Schutzmassnahmen nach dem Opportunitätsprinzip vorgehen muss. Es kommt meistens aus der Politik, nicht nur aus der Sachlage oder der Objektivität. Im Idealfall aber kommt aus dem realen, objektiven Problem via die Politik eine Dynamik rein. Aber meistens ist es sehr politisch, ob man etwas unternimmt oder nicht. Nur schon die Erkenntnis auf der Stufe der Gemeinde: Man muss und will etwas machen. Wir wollen Mittel aufwenden. Wir haben ein Problem. Auch diese Erkenntnis ist sehr wichtig. Natürlich kann dort eine Risikokarte helfen, aber auch wenn man… Wenn ich mir es richtig vorstelle, wenn du eine Karte mit Unsicherheiten oder Sicherheiten hast. Sprich eine hohe Wahrscheinlichkeit für die Betroffenheit eines Gebäudes. Wenn du das dann noch flächig hast für viele Gebäude… Es geht schon ein wenig in diese Richtung, dass man Hotspots ableiten könnte. Das ist eine wichtige Grundlage. Ich bin einfach der Meinung, dass diese Tools, welche wir haben, also in der Praxis reicht es - ohne deine Arbeit in Frage stellen zu wollen. Und es darf nicht kompliziertere Tools geben, denn das ist kontraproduktiv. Es bremst die ganzen Verstrebungen aus. Dann sage ich lieber irgendwo… Es geht auch dort wieder darum, wo die Sicherheit und Unsicherheit ist. Dann mache ich lieber eine Massnahme in den nächsten 5-10 Jahren, welche einen gewissen Teil der Stadt oder des Dorfes schützt. Auch wenn der Schutz etwas zu hoch oder zu gering dimensioniert ist, bringt er einfach eine wesentliche Verbesserung. Anstatt dass man in 50 Jahren noch gar nichts verbessert hat, weil man dieses, jenes und noch mehr Planungsinstrumente hat. Das ist dort… Ich bin der Meinung, dass wir uns auf einem Niveau befinden, auf welchem wir ausreichende Grundlagen haben. Die Hauptlücke oder das, was ich vor allem pushe, sind im Hochwasserschutzbereich die flächigen Karten der Fliessgeschwindigkeiten und der Fliesstiefen. Das ist das Produkt, das in der Praxis hilft. Auch wenn das schwarz-weiss ist und 1.25m Überflutung auf den cm genau angibt. Aber dann habe ich zumindest einen Wert, an dem ich mich orientieren kann und die Unsicherheit dieser Wasserspiegellage kann ich selbst abschätzen. Und ich weiss, über ein Gebiet wurde die gleiche Methodik angewandt. Wenn ich am einen Ort 1.25m habe und an einem anderen Ort habe ich 67cm und ich weiss die Methodik dahinter ist die gleiche, dann kann ich das einordnen und kann dort… Mir ist dort absolut bewusst, dass diese Wasserspiegellage nicht jene ist, welche am Tag X beim 100-jährlichen Ereignis auftreten wird. Aber es zeigt eine Richtung an.
JB	*Dann wäre ich mit meinen Fragen durch. Gibt es noch etwas bezüglich der Gefahrenkarten, deiner Arbeit oder zu Murgängen, was du noch erläutern möchtest?* *Alright, those were all the questions I prepared. Are there other inputs on hazard maps, your work and regarding debris flows that you would like to mention?*
E4	Nein, ich finde es spannend. So wie ich es verstanden habe, erstellst du auch Karten. Im Grafischen würde es mich noch interessieren. Ich bin auch sehr Karten-affin. Es gibt sehr spannende Kartendarstellungen. Ich finde es wichtig und richtig, dass man etwas in diese Richtung macht. In der Praxis müssen wir einfach Abstriche machen, was die Komplexität anbelangt, aber das ist auch nicht schlimm. Ich denke, die Anwendung und die Forschung sind dort zwei Paar Schuhe. Sie müssen zueinander passen, müssen sich

	ergänzen, aber es kann nicht der Anspruch sein, dass man in der Praxis das Forschungsniveau erreicht. Das ist schlicht nicht handelbar.

E. Transcript of Expert Interview E5

Interviewer:	**Interviewee:**	**Role:**	**Date:**
Jana Bracher	E5	Practitioner	22.09.21

This interview was conducted in Swiss German and translated to German during the transcription which resulted in the rephrasing of certain statements. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability.

JB	*Für dich ist es in Ordnung, wenn ich das Interview aufzeichne?* *Is it okay for you, if I record the interview?*
E5	Jawohl, das ist gut.
JB	*Dann starten wir mit den Fragen. Erkläre bitte kurz, was dein daily-business in der Arbeit mit Naturgefahren ist.* *Then, let's begin with the questions. Please explain everyday tasks of your position.*
E5	Genau, also ich beziehe mich hier auf die Arbeit hier in D.. Aktuell. Ja grundsätzlich sind es eigentlich 3 Bereiche, die mir in den Sinn kommen. Das eine ist grundsätzlich der Unterhalt oder die Unterhaltsplanung forstlicher Schutzbauten. Das läuft zwar über den Forstdienst auch hier, wo der Förster auch stark involviert ist. Aber es ist dann ein Thema, welches in der Gemeinde… das jährliche oder 5-jährliche Unterhaltsprojekt, welches man für die Subventionen eingeben muss, das läuft über uns, hier auf der Bauverwaltung. Und wenn hier Schäden sind, gibt es auch ein Kreditgebrauch, so läuft das dann über uns. Das ist der Naturgefahrenbereich. Der zweite Bereich ist die Zonenplanung also die Raumplanung selber. Wenn Schutzbauten ausgeführt wurden, wenn es Anpassungen in der Gefahrenkarte gab, welche man dann in den Zonenplan Naturgefahren übertragen wird, sodass das dann auch abgebildet sind und natürlich auch sonst im Zusammenhang mit Zonenplanung werden Naturgefahren nach Möglichkeiten berücksichtigt. Und dann im Rahmen der Baubewilligungsverfahren, seien es Projektvoranfragen oder eingehende Baugesuche, bei welchen man auch prüfen muss: Ist das in einem Gebiet, in dem Naturgefahrenprozesse existieren nach der Gefahrenkarte, ist es betroffen oder nicht betroffen? Von welchem Prozess? Und entsprechend leiten wir das dann an die Fachstelle zur Beurteilung weiter. Sie geben ihre Stellungnahme ab im Sinne von: es geht oder es geht nicht, oder sie formulieren Auflagen. Das ist das, was ich gerade so grob…
JB	*Wie wichtig ist die Gefahrenkarte als Instrument in deiner Arbeit?* *How important is the hazard map as an instrument in your work?*
E5	Das ist natürlich sehr wichtig, diesbezüglich. Weil einerseits ist es eine einfache Visualisierung von diesen Prozessen und diesen Gefahrengebieten, Gefahrenstufen – rot, blau, gelb, Restgefährdung oder gar nichts. Und ja, da kann auf einen Griff, sei es auf Papier oder digital, kann man so, ja, einem Kunden in Anführungs- und Schlusszeichen, einem Baugesuchsteller oder jemand, der eine Frage stellt, kann man das kundtun. Sie

	ist unterdessen auch für jedermann zugänglich. Das dient sehr der Transparenz und es ist eine einfache Methode aus meiner Sicht zur grafischen Darstellung, um schnell zu entscheiden.
JB	*Welche Entscheidungen triffst du mit den Gefahrenkarten? Spezifisch für Murgänge?* *What are decisions that you make with hazard maps?*
E5	Ja, Entscheidungen, welche wir fällen müssen, ist eben einerseits, welche Fachstelle wir bedienen müssen mit einem entsprechende Baugesuch / Bauvorhaben. Und wir können auch schon in der Beratung, wenn jemand fragen kommt: Ich möchte dies machen an diesem Standort. Dann können wir schon sagen: Achtung, rotes Gefahrengebiet… wahrscheinlich solltet ihr die Finger davon lassen. Blau – vielleicht schon frühzeitig mit der Fachstelle Kontakt aufnehmen und das entsprechende Projekt schon einmal mit ihnen diskutieren, allenfalls schon Optimierungen bei der Projektplanung machen. Das ist sicher ein wichtiger Punkt. Und ja, im Zusammenhang mit der Raumplanung, wo wir auch schon sehen, hier ist ein Gefahrengebiet, da hat es keinen Wert, wenn wir überhaupt an eine Einzonung denken. Zum Beispiel, das sind Sachen, wobei die Gefahrenkarte ein gutes Instrument ist.
JB	*Da sind also auch eine Vielzahl von Akteuren involviert in solchen Entscheidungsprozessen?* *So there are a large number of actors which are involved in a decision-making process?*
E5	Das kommt ein wenig darauf an. Wenn jemand eine Anfrage stellt oder wenn ein Bauherr / Projektverfasser oder so kommt, dann ist es im Normalfall, also sie kommen zu uns und wir geben eine Antwort und dann entscheiden sie selber: machen sie weiter, oder machen sie es nicht. Gehen sie an die Fachstelle oder gehen sie nicht an die Fachstelle. Wenn es dann darum geht, Zonenplanung ein Gebiet einzuzonen, umzuzonen oder auszuzonen, dann ist es das zuständige Organ, welches darüber entscheiden muss. Es kommt ein wenig auf den Planungsschritt an. Wenn es im Rahmen der Schlussfassung ist, wenn es die Vordiskussion ist, ist es auch oft im kleineren Rahmen. Dass man es diskutiert und es dann schon verwirft bevor es zum Gemeinderatsprojekt wird oder sogar an die Gemeindeversammlung kommt.
JB	*Dann wird die Gefahrenkarte sehr früh im Prozess beigezogen?* *So the hazard map is part of the process early on?*
E5	Ja, das ist so. Im Normalfall ist es auch so, dass man das Baugesetz nicht mehr vergisst. Das ist vielleicht vor 10-15 Jahren anders gewesen, als man die Gefahrenkarte noch auf Papier war und man hat die noch nicht so genutzt. Und heutzutage ist es etwas vom ersten, was man macht: die Gefahrenkarte hervornehmen und schauen, wo man liegt.
JB	*Vielleicht noch die direkte Frage, welchen Zweck erfüllt die Gefahrenkarte für euch in eurer Arbeit? Ist eine Entscheidungsgrundlage, eine Orientierungshilfe?* *Just to ask directly, what is the purpose of the hazard map for you?*
E5	Ja beides. Beides, oder. Einerseits ist eine Orientierungshilfe, damit man einfach etwas abschätzen kann und ein entsprechendes Projekt darauf ausrichten kann. Dass man sagen kann, das muss man berücksichtigen, vergesst es nicht oder nehmt Kontakt mit der Fachstelle auf. Oder manchmal ist es auch so, dass man sagt, ja, rote Gefahrenzone, das muss man gar nicht diskutieren, da muss man gar nicht zur Fachstelle, sondern kann es gleich vergessen. Wobei natürlich die Entscheidungskompetenz in dem Sinne in seltenen Fällen bei uns liegt. Es sind nicht wir, die sagen können, vergesst es. Wenn der Bauherr kommt und sagt, er macht es trotzdem, dann ist es doch die Fachstelle, die sagt nein. Das sind nicht wir als Baubehörde zum Beispiel, die darüber entscheiden. Wir übernehmen dann die Haltung der Fachstelle, ihre Stellungnahme in den Bauentscheid zum Beispiel übernehmen. Dann ist rot, nein gleich nein. Da dürfen wir nicht etwas anderes schreiben. Darum als Gemeinde für Gemeindeeigene Projekte ist es klar. Da sind wir daran und sagen: bei einem Strassenabschnitt oder einen Leitungsverlauf ziehen wir besser nicht durch ein Gefahrengebiet, sondern ziehen sie ausserhalb durch. Dann ist es eigentlich schon eine Entscheidungsgrundlage.

JB	*Fördert die Gefahrenkarte in der heutigen Form deine Entscheidungen dadurch, dass sie eher grob ist mit 3-4 Stufen oder gibt es auch Herausforderungen, weil die Klassifizierung so grob ist?* *Do hazard maps in their form today with only three levels of hazard support your decisions or are there challenges because the classification is so broad?*
E5	Grundsätzlich bin ich der Meinung, dass es der Sache dient. Es ist klar, das ist auch etwas, was man gut merkt. Die Leute wissen das. Es gibt fast niemanden mehr, mit dem man redet und sagt, was ist eine rote Gefahrenzone, was bedeutet das. Das wissen die Leute eigentlich. Es ist auch das Ampelsystem ein wenig mit den Farben der Gefährdung. Das dient. Wo es schwieriger wird, ist wo die Intensität, wie soll ich sagen, so die Grenzbereiche: sind wir noch gelb oder ist es schon blau, ist es ein intensives blau oder ist es eher ein schwaches blau. Wobei auch hier habe ich die Erfahrung gemacht, da muss man mit Fachleuten sprechen, die erklären das und leiten die Massnahmen daraus ab. Dann ist das auch klar. Im Grundsatz ein sehr gutes Instrument, sehr einfach, ein sehr transparentes System und es ist verständlich für die Leute.
JB	*Werden Entscheidungsträger, ihr auf der Gemeinde oder auch Bauherren, in der Verwendung der Gefahrenkarte geschult oder wird sie einfach öffentlich zur Verfügung gestellt?* *Are decision-makers, in the municipality or house builders trained in the usage of the hazard map or is it just presented openly?*
E5	Ja, wir haben natürlich in der Ausbildung, je nach dem welche Ausbildung man macht. Es gibt keine spezifische Naturgefahren-Ausbildung. In der Bauverwalterausbildung sind die Naturgefahren ein Bereich. Das ist aber sehr dürftig, rein von der Breite her. Aber das reicht. Die Gefahrenkarte ist eines, zusammen mit den entsprechenden Gefahrenprozesse und den Gefahrenstufen und so, ist ein Thema. Und ich glaube das ist ausreichend. Und wenn man möchte, und das ist bei uns im Kanton und auch bei uns auf der Gemeinde so, man bekommt von der Fachstelle die nötige Information. Ich glaube es bringt nicht viel, weil es sofort sehr komplex wird, wenn man hier auf der Gemeinde mehr wissen müsste oder sollte. Weil Bescheid geben ist immer noch schwieriger. Das einerseits zu verstehen ist das eine, aber es zu erklären ist etwas anderes. Aber da sind eigentlich auch die Fachstellen sehr gut anerkannt und sie helfen, sind sehr kooperativ. Man hat dort immer Unterstützung, wenn wir es brauchen. Wenn jemand eine Frage hat oder wenn etwas geklärt werden muss können wir per Mail oder Telefon oder sogar in einer Begehung / Besprechung, kann man dem begegnen.
JB	*Dann ist die Fachstelle auch gegenüber der Öffentlichkeit die Anlaufstelle?* *So the speical department is also the point of contact for the public?*
E5	Ja. Also da gibt es schon Leute, welche direkt an die Abteilung Naturgefahren gehen, wenn es um Lawinen oder Rutschungen, Steinschläge geht oder auch Wassergefahren. Das wissen sie aus anderen Bauvorhaben oder weil es einmal berichtet wurde oder weil sie sogar die Leute noch kennen. Das sind vielleicht auch Themen. Gerade mit den grossen Hochwasserschutzprojekte, welche im Oberland sehr präsent sind. Oder Lawinenverbauungen oder so, das ist Themen, welche immer in der Presse ist. Wenn jetzt irgendwo… sagen wir A. hat 1.2 Mio beschlossen für eine Lawinenverbauung und dann gibt es einen Bericht in der Zeitung und dann ist N. H. von der Abteilung Naturgefahren, gibt da auch noch ein Interview. Dann ist das immer ein Thema. Und dann die Gefahrenkarte: wieso macht man das? Das ist immer auch ein Thema.
JB	*In welcher Form verwendest du die Gefahrenkarten? In gedruckter oder digitaler Form?* *How do you use hazard maps? In digital or in printed form?*
E5	Beides. Jetzt inzwischen ist sie auch immer sehr aktuell, wir können sie überschneiden mit anderen Informationen. Das ist sehr nützlich. Mit dem Zonenplan oder mit Grundeigentum, der amtlichen Vermessung und mit anderen Themen. Das ist natürlich sehr praktisch. Auf der anderen Seite ist es gut, wenn man die Papierform hat. Es ist schon nur die Archivierung, es ist so… was gilt? Digital, ja man kann das

	schon digital genehmigen, was auch ein Thema wird jetzt, dass man diese Sachen auch digital rechtsverbindlich macht. Aber es wird eine Herausforderungen sein immer das aktuelle und das rechtsverbindliche Dokument digital immer am richtigen Ort zu haben. In Papierform ist es etwas einfacher. Und wenn man raus geht, dann liegt es auch der Hand. Dann nimmt man einen Ausschnitt des Zonenplans, respektive des Gefahrenkarte mit, oder man nimmt den ganzen Plan mit je nach dem. Dann ist es auch einfach etwas darzustellen, zu zeigen.
JB	*Zukünftige Naturgefahren sind mit hohen Unsicherheiten verbunden. Die Gefahrenkarten sind in nur 3 Kategorien aufgeteilt: Wenn ihr mit der Gefahrenkarte arbeitet, ist man sich dann dieser Unsicherheit bewusst? Future natural hazards are influenced with a lot of uncertainty and the hazard maps only distinguish three levels of hazard: When you are working with the hazard maps, are you aware of these uncertainties?*
E5	Ja, das ist noch schwierig. Die Erstellung von Gefahrenkarten, die Erarbeitung, der Prozess dahinter, sehen die meisten Leute, welche damit arbeiten, zu wenig. Oder, weil ja… was hat man, ich sage es ein wenig provokativ… wenn die Baumlänge 30m angenommen wird bei einem Murgang, oder die Verklausungsgefahr, oder nimmt man 15m an. Das ist wahrscheinlich ein relativ wesentlicher Aspekt bei der Berechnung einen Brückendurchlasses. Und ja irgendwie geht man davon aus, sie haben das schon richtig gerechnet. Und das Resultat, dass das eine gewisse Unschärfe hat, ist klar. Aber man nimmt es vielleicht ein wenig zu genau. Sagen wir es so. Und dort ist es ein wenig die Schwierigkeit, wenn man es noch auflösen möchte, ist es noch so nachvollziehbar für die Leute, so logisch und so klar, was es bedeutet? Und ich habe im Moment das Gefühl, dass das System, wie wir es jetzt haben, ist nicht so schlecht. Es ist dann wieder die Frage bei der Erarbeitung oder… wie weit geht man in die Details und definiert alle Parameter aufs Detail. Oder müsste man eine gewisse Unschärfe lassen und dann fällt es vielleicht einmal statt ins blaue doch noch ins rote. Also nicht, dass man da noch eine riesige Sicherheit einbauen muss, aber nicht zu speziell und zu konkret alles auseinandernehmen, damit man dann sagen kann: es ist 0.99 und nicht 1. Ja nein 0.9 reicht aus, denke ich. Und der Rest ist dann 1, als Beispiel. Ich habe diesbezüglich das Gefühl, mit einer genaueren Berechnung, respektive wenn man etwas noch genauer deklariert, entfernt man auch ein wenig die Eigenverantwortung. Ich sage jetzt mal die Oberflächenabflüsse, es ist eigentlich verrückt, dass man das den Leuten sagen muss: Achtung, wenn ihr eine Tiefgarage macht oder den Lichtschacht im Keller unter Terrain nehmt, könnte Wasser reinlaufen. Eigentlich müsste das aus der Erfahrung der Planer klar sein, Architekten und auch der Bauherr sollte ich selber etwas überlegen und nicht zur Gemeinde kommen und sagen: Das habt ihr mir bewilligt, ihr habt nicht besser geschaut. Also irgendwo muss man vermeiden, dass man zu viel Verantwortung auf die Fachstelle und die Bewilligungsbehörde schaufelt. Das könnte kontraproduktiv werden. Und ja… was ich da auch denke, ist, manchmal hat man besser etwas im zu starken Gefahrengebiet als im zu schwachen. Der Perimeter ist vielleicht einmal besser zu gross als zu klein. Man kann das jetzt vielleicht sagen aus der Sicht, das ist jetzt wieder eher L.. Man hat dort Lawinenverbauungen gemacht im M. für ca. 35 Mio., welche das Dorf W. schützen und insbesondere auch der touristisch genutzte Bereich und eigentlich praktisch bis zu Dorfstrasse, wo man in den letzten 70-80 Jahren Lawinenereignisse gehabt bis zur Dorfstrasse. Mit der Verbauung jetzt logischerweise nicht mehr. Die Verbauung hat eine Lebensdauer von geschätzten 80 Jahren. Man hat Ende 70ern angefangen zu bauen, also die ältesten Verbauungen sind inzwischen auch schon bald 50-jährig. Also hat jemand noch die Finanzen, diese Verbauungen in 30-40 Jahren zu erneuern. Und wenn nicht, was machen wir, wenn wir die Gefahrenkarte anpassen und das als Bauzone ausscheiden, was bis jetzt rote Zone war und dann sind diese total im Ding drin. Dort habe ich das Gefühl, dort müsste man wahrscheinlich noch ein bisschen einen Schritt weiter gehen. Das ist jetzt meine persönliche Meinung, da kann ich nicht für die Gemeindebehörde reden. Aber weil… du kannst fast wie die organisatorischen Massnahmen, man kann sagen, man hat Massnahmen ergriffen, sodass man das Risiko, dass etwas passiert, minimiert. Aber das ist nicht unbedingt nachhaltig, sodass man gleich

	die Gefahrenkarte anpasst. Also, zumindest dort, wo es nicht überbaut ist, das ist das, was ich… jetzt bin ich schon fast bei der letzten Frage… eine Differenzierung, oder eine differenzierte Anwendung im Bereich von bereits überbauten, man muss sagen: hier steht schon ein Haus. Da kann man sagen, doch eine gewisse Hürde muss man hier vielleicht abbauen. Im Sinn von: wenn er eine Baubewilligung beantrage für ein zusätzliches Fenster, eine Aufstockung, ein Geschoss mehr im Extremfall. Dann muss man sagen, der ist im geschützten Bereich, da haben wir keine Auflagen bezüglich der Naturgefahren. Aber wenn das eine Parzelle ist, welche bis jetzt nicht überbaut ist, bin ich der Meinung, müsste man sagen: Schaut, man hat zwar Massnahmen ergriffen, aber man hat die Massnahmen nicht ergriffen, dass ihr jetzt hier bauen könnt. Sondern man die Massnahmen ergriffen, dass diejenigen, welche im Gefahrengebiet jetzt drin sind, das sie geschützt sind. Es geht ja um die Sachwerte, welche man jetzt hat und nicht, dass man zusätzliche Sachwerte ins geschützte Gebiet baut. Und dann in 50 Jahren sagen sie, das wären ja schöne Idioten. Wir können das ja gar nicht mehr handeln. Mit den Überlastfällen sind es ähnliche Probleme. Mit den Murgängen M. oder auch G.… neue Berechnungsmethoden und dann hast du plötzlich, das was zurückgestuft wurde ins blaue oder gelbe, kommt plötzlich wieder ins blaue oder sogar ins rote. Und das sind natürlich die Unsicherheiten, ich denke dort braucht es eine Art von einem zweistufigen Verfahren, dass man sagt: Das ist Anwendungsbereich überbaut – Besitzstand und Anwendungsbereich neu. Dort muss man viel schärfer sein. Das schwebt mir vor.
JB	*Du hast vorhin die Herausforderungen in Grenzgebieten erwähnt. Wird dort eine Art von Unsicherheit kommuniziert und es wird gesagt, die Einteilung ist grob oder geht man nach dem schwarz-weiss Schema? Denkst du die räumlichen Unsicherheiten, welche bestehen, sollten in Karten dargestellt werden?* *You mentioned before the challenges in the border regions. Are uncertainties in those regions communicated and you tell people that the classification is broad or does one follow the black and white scheme? Do you think that spatial uncertainties tied to those phenomena should be displayed?*
E5	Man probiert es den Leuten schon zu erklären. Aber es gelingt, einerseits gelingt es nicht jedes Mal und andererseits denkt man auch nicht immer daran, man vergisst es auch. Es ist schon so. Primär im Vordergrund steht rot-blau-gelb-gelb/weiss oder keine. Das ist die ultimative Antwort, die sie suchen: bin ich drin oder bin ich nicht drin. Dass man dann sagt: schaut, ein gewisses Restrisiko in Anführungs- und Schlusszeichen oder ein Risiko im Zusammenhang mit Veränderungen in den Verhältnissen, sei es von den Prozessen her oder auch von der Berechnungsmethode oder was auch immer. Das Ereignis gilt und man ist besser auf der sicheren Seite als auf der unsicheren. Das ist das, was ich vorhin gesagt habe. Die Kosten bei einem Neubau, wenn man etwas neu realisiert oder etwas aufgleist, sei es eine Erschliessungsleitung oder ein Strasse oder so, die sind natürlich im Normalfall deutlich tiefer dort das zu berücksichtigen oder eben nicht zu realisieren und eine Alternative zu machen, als wenn man es realisiert hat und dann muss man plötzlich Massnahmen ergreifen wegen eines Ereignisses, welches kam, oder man hat sogar den Schaden darauf. Und ja… das sehe ich so.
JB	*Und wenn man die Unsicherheiten grafisch darstellen würde. Würde die Darstellung der Unsicherheiten den Entscheidungsprozess komplexer machen oder würde es auch helfen, indem man in der Karte noch mehr Informationen hat?* *If we were to display uncertainty grafically. Would it make the decision-making process more complex or would it also help because of the higher information content in the map?*
E5	Also du würdest zum Beispiel vorsehen, oder als möglich Lösung sehen, dass es irgendwie ein dunkel-blau gibt, oder ein hell-rot (*JB: Genau, eine kontinuierliche Darstellung.*) Ja, wahrscheinlich würde es dem Verständnis der Leute nicht im Weg stehen, ich sage es so. Der Normale würde das wahrscheinlich lesen und sagen, es ist ja logisch: zwischen rot und blau gibt es irgendwo einen Übergang, der hat eine etwas andere

	Farbe, das ist etwas zwischen rot und blau. Und dann gibt es da dunkelrote und das, ja ich weiss auch nicht grün zwischen gelb und blau. Einfach so etwas. Auf der anderen Seite würde wahrscheinlich die Schwierigkeit losgehen: Wo hat man effektiv noch welche Massnahme? Ist man im hellroten oder dunkelblauen… ist das schon Bauverbot oder nicht? Und dann wäre man eigentlich fast schon wieder beim, also das ist wieder spontan, beim Lösungsansatz von vorhin. Man müsste konsequenter sagen, dort wo noch gar nichts ist und der Gefahrenprozess schwierig abzuschätzen ist und die Unsicherheit gross ist, da müsste man sagen, alles was eine Frage hat ist gesperrt – also Bauverbot. Nur für standortgebundene und absolut dringend nötige Sache, die man nicht anders machen kann. Leitungsbau, wo man bei Gottes Namen dort durch muss oder eine Strasse, die man einfach dort machen muss, was nicht anders geht. Aber irgendwo, wo bereits eine Scheune oder so… das baut mal dann nicht dort, weil dort ein gewisser Gefahrenprozess betroffen ist. Auch wenn dieser nur dunkelgelb oder hellblau oder dunkelblau oder was auch immer. Dann wäre es wahrscheinlich einfacher. Weil sonst geht die Diskussion los: das ist jetzt im dunkelblauen, das ist im hellblauen noch. Das machen oder jenes. Welche Massnahmen muss man ergreifen? Dann würde es fast wieder für das Gegenteil genutzt. Man würde Abstufungen machen, welche man eigentlich auflösen wollte. Einfach noch mehr Abstufungen, da sage ich: ja gut bei hellblau muss das Fenster 1m ab Boden haben und bei dunkelblau muss es 1.5m ab Boden haben. Wegen der Intensität vom Ereignis. Ob das dem entspricht, was du dir vorstellst, weiss ich nicht.
JB	*Sprich es würde das Verständnis vom Prozess bei den Leuten vereinfachen, aber politisch würde es komplizierter. Wärst du damit einverstanden?* *So it would help the process understanding for people but on a political level it would become more complicated. Would agree with this statement?*
E5	Ja, in der Umsetzung würde es schwierig werden. Und ja, genau. Das ist sicher so. Und das Verständnis der Leute… ja, das ist noch schwierig zu sagen, weiss du, wie das Verständnis… ob man das wirklich in der Gefahrenkarte in der Darstellung abholt oder ob man das auf eine andere Art… Ich sage jetzt ein Murgang- oder Steinschlagmodell, das ist etwas, wo viele Leute keine Ahnung haben wie das läuft. Da könnte man sagen, man könnte Öffentlichkeit, wenn eine neue Gefahrenkarte kommt, muss man einen Informationsanlass oder so etwas einbauen oder auch mal in einem Pressebericht. Dass da effektiv Modelle dahinter stecken und das wie beim Wetter ist. Die Wetterprognose für das nächste Wochenende. Die Strömung von Süden ist anders als sie am Schluss ankommt. (JB: *Ja genau darum geht es, einfach über andere Zeithorizonte.)* Und das ist vielleicht auch noch ein Aspekt oder. Man hört… das ist etwas, das noch viel verbreitet ist in den Köpfen der Leute: das 300-jährliche. «In 300 Jahren sind wir ja nicht mehr hier.» (*lacht*) Das wird relativ stark, da merkt man, dass das Verständnis von der Statistik bei vielen oder bei den meisten Leuten nicht vorhanden ist. Weil das ist nicht ganz mit unserem logischen Denken vereinbar, die Statistik, Wahrscheinlichkeitsrechnung. Das verstehen viele nicht. Und dann haben sie das Gefühl: Ja das 100-jährliche.. in den letzten 100 Jahren ist ja nie so etwas passiert. Also stimmt es nicht. Ja doch, 100-jährlich heisst einfach die Wahrscheinlichkeit ist, dass es alle 100 Jahre passiert. Es kann 1'000 Jahre lang nicht passieren und dann passiert es 10 Jahre nacheinander jedes Jahr theoretisch. Und das, das bringt man nicht in die Köpfe. Und dort redet man vielleicht zu stark von dieser Wiederkehrdauer und der Ereigniswahrscheinlichkeit hinter dem Prozess und sagt, das ist ein Modell und je nach dem wie das Modell die Ausgangsgrössen definiert, geht die Lawine 100m weiter runter als nicht und da hat niemand recht. Es ist beides vielleicht richtig, wir wissen es nicht. Und das könnte vielleicht dazu beitragen, dass man dort ein wenig noch mehr abgewinnt in der Bevölkerung. Dass die Bevölkerung eine gewisse Unsicherheit, auch wenn es rot-blau-gelb ist, dass gewisse Unsicherheit trotz allem besteht.

JB	*Mit meinen vorbereiteten Fragen bin ich dann durch. Gibt es noch etwas bezüglich der Gefahrenkarten, deiner Arbeit oder zu Murgängen, was du noch erläutern möchtest?* *Alright, those were all the prepared questions. Are there other inputs on hazard maps, your work and regarding debris flows that you would like to mention?*
E5	Nein, ich glaube das ist schon das. Und ich denke, du hast das richtige angesprochen mit den Veränderungen. Es ist nicht immer gleich wie in den letzten 50 Jahren. Die Witterung ist nicht immer gleich, die Entwicklung, die Bebauung, die Nutzung von der Umgebung und alles zusammen. Die Ganze, ja, ich sage mal Rutschungen oder so, das Material, welches zur Verfügung steht. Das hat alles einen Einfluss. Und da hat man schon eine gewisse Unsicherheit und das ist schon ein wichtiger Aspekt. Es geht nicht nur um Klimawandel in Anführungs- und Schlusszeichen. Es geht grundsätzlich darum, dass man sich bewusst ist, die Natur ist kein Rechenschieber. Es ist einfach… manchmal passiert etwas, das hätte man gar nicht gedacht. Genau. Sonst habe ich eigentlich nichts direktes. Das ist gut.

F. Transcript of Expert Interview E6

Interviewer:	Interviewee:	Role:	Date:
Jana Bracher	E6	Practitioner	13.10.2021

This interview was conducted in Swiss German and translated to German during the transcription which resulted in the rephrasing of certain statements. Filler words, repetitions, incomplete sentences (indicated as …), statements, which did not relate to the topic (indicated as (…)), as well as confirmative comments from the interviewer are excluded from the transcript to enable a better readability.

JB	*Ist es für dich in Ordnung, wenn das Interview aufgenommen wird?* *Is it okay for you if the interview is recorded?*
E6	Ja, das ist in Ordnung.
JB	*Erkläre bitte kurz, was dein daily-business in der Arbeit mit Naturgefahren ist.* *Please explain some of the everyday tasks of your position regarding natural hazards.*
E6	Mein daily-business, das ist ursprünglich… Am besten mache ich es chronologisch. Ursprünglich bin ich in dieser Zeit hinzugekommen… also ich habe 2003 hier angefangen zu arbeiten und habe eigentlich genau diese Zeit getroffen, in der die Gefahrenkarten im grossen Stil gemacht wurden. Das heisst, wir haben relativ viele Berner Oberländer Gefahrenkarten gemacht, nicht nur dort aber viele. Und zu diesem Zeitpunkt habe ich vorwiegend im Bereich Murgang gearbeitet. Natürlich auch sonstige Wassergefahren und Überflutungen, aber vor allem Murgänge. Weil im Gebirge sind Bäche in der Regel sofort Murgang-lastig. Das waren jeweils 2-3 Gefahrenkarten, welche parallel erstellt wurde. Das ging etwa ein Jahr, also ein relativ ruhiger Job. Als dies abgeschlossen war, kamen dann immer mehr die Spezialfälle. Das sind Ereignisse wie jenes, welches ich im S. präsentiert habe. Einfach aussergewöhnliche Ereignisse auch an anderen Orten, welche man zu interpretieren, verstehen und analysieren begonnen hat. Und man hat gemerkt, dass es über die Gefahrenkarten hinaus geht. Das kann man nicht mehr mit diesem Blickwinkel betrachten. Es wurden immer mehr die Spezialfälle. Einerseits blieben das die Murgänge. Dann kamen mit dem 2011er Ereignis, welches nicht nur im S. aber auch an anderen Orten im Berner Oberland enorm viele Murgänge aus Permafrost

	Gebieten gebracht hat, mehr Permafrost-Studien und Naturgefahren im weitesten Sinn, eher überblicksmässig. Was ich sonst noch parallel mache seit ca. 2011, ist die Gletscherseegeschichte, welche bis heute anhält. Da bin ich heute gerade daran, einen Bericht zu schreiben. Und was wir häufig nebenher haben, sind Objektschutzgutachten. Wenn jemand einen Umbau oder ein Haus plant, dass man dort Schutzmassnahmen dimensioniert. Wie hoch müssen die Fenster sein? Wenn man sie tiefer möchte, wie viel Druck müssen sie standhalten? Da modelliert oder rechnet man es dann detailliert für ein Gebäude. Das mache ich am häufigsten. Unterdessen behandle ich mehr Lawinen als Murgänge abgesehen von den Spezialprojekten, welche immer weiter gingen. Das ist es im weitersten Sinn. Und manchmal mache ich noch einige Hangmuren und gewisse Rutschbeurteilungen, aber das ist eher untergeordnet.
JB	*Du sagtest, dass du nahe an der Erstellung der Gefahrenkarte gearbeitet hast. Beschreibe bitte das Vorgehen, wenn Gefahrenkarten erstellt werden. Evtl. gerade bei den ersten Versionen.* *You said that you closely worked in the generation of the hazard maps. Please explain the process of creating the hazard maps. Maybe especially for the first versions.*
E6	Die Erstellung ist eigentlich immer ähnlich bis heute, ausser dass… Es ist vor allem die Grundlagenbeschaffung und deren Studium. Das beinhaltet insbesondere die Gefahrenhinweiskarte, welche ich in der Präsentation gezeigt habe. Der Ereigniskataster, um zu sehen, wo bereits etwas passiert ist. Die Gefahrenhinweiskarten, ich weiss nicht, ob du dich daran erinnerst, die werden nur modelliert und zu sehen, wenn der «Supergau» grob für alle Prozesse gerechnet wird, wo muss ich gar nicht nachsehen, wo müsste ich trotzdem noch hinschauen, auch wenn es nicht so scheint, weil dort vielleicht etwas unerwartetes auftritt. Und dann der Ereigniskataster… Man rechnet dann die Hydrologie: Welche Abflüsse können stattfinden? Wenn das alles durchgerechnet wurde, hat man Anhaltspunkte bezüglich der Grösser der Bäche und der Menge an Wasser, die kommen kann. Dann geht man ins Feld und sieht zum Beispiel wie viel Geschiebe vorhanden ist. Und man nimmt an den engsten Stellen die Querprofile auf, um zu sehen, ob die berechneten Wassermengen, das Geschiebe oder das Holz unter einer Brücke durch passen oder nicht. Dann geht man zurück ins Büro und rechnet es durch und vergleicht die Querprofile mit dem Abfluss. In der Regel besprechen wir das dann in einer sogenannten Szenarien-Sitzung mit dem Kanton und der zuständigen Person. Da kann man seine Überlegungen erklären. Und sie geben vielleicht auch noch Inputs, welche aus ihrer Sicht wichtig sind. Dann gibt es eine Sitzung, um die Szenarien der Gemeinde zu präsentieren. So wissen auch sie, in welche Richtung es geht, und sie können ihre Meinung abgeben. Und dann beginnt man, die Gefahrenkarte zu zeichnen. Oft verwendet man Modellierungen, nicht immer, vor allem früher war es noch nicht möglich, dies zu modellieren. Mittlerweile macht man an Orten, wo die Topografie klar ist, sogar keine Modellierungen mehr. An anderen Orten, wo die Lage nicht klar ist, macht man sie noch. Dann gibt es eine Entwurfsbesprechung mit dem Kanton, wie bei den Szenarien. Und man bespricht die Resultate auch mit der Gemeinde an einer Sitzung. Sie bekommen dann alle Entwürfe und können diese ansehen und der Kanton und die Gemeinde geben dann eine Rückmeldung. Dann macht man das Finish. Das war in etwa der Ablauf. Heute ist es grundsätzlich noch gleich, ausser, dass die Ansprüche unglaublich gestiegen sind, das ist unglaublich. Auch der Unterschied in der Beurteilungsflughöhe der Gefahrenkarte zwischen dem Anfang und heute. Mittlerweile muss man für jede einzelne Prozessquelle eigene Intensitätskarten rechnen. Man muss alles durchrechnen und nicht… Häufig hat man Übertragungsüberlegungen gemacht z.B., wenn sich dieses Gerinn ähnlich verhält wie diese fünf nebenan, dann machen wir Übertragungsüberlegungen. Es werden viel mehr Details gefragt, alles muss erklärt sein. Und das andere, das sich geändert hat, ist, dass man häufig noch Zusätze rechnet. Zum Beispiel bei Abflüssen, weil man weiss, dass mehr intensive Regenfälle zu erwarten sind.

JB	*Und kommen die höheren Ansprüche eher vom Fachlichen oder kommen sie von denjenigen, z.B. die Gemeinden und Kantone, die die Gefahrenkarten in Auftrag geben und sie schlussendlich brauchen?* *Do those higher standards come from the specialist's side or from those who commissioned the hazard maps such as municipalities or cantons?*
E6	Ich glaube es kommt von beidem. In erster Linie sicher der Kanton. Sie wollen sich doppelt und dreifach absichern. Beziehungsweise absichern ist vielleicht das falsche Wort. Sie sehen natürlich auch, wofür sie sie (die Gefahrenkarten) sonst noch brauchen. Wenn man beispielsweise die Hydrologie rechnet und sie später noch anderorts ein Verbauungsprojekt haben im Wasserbau, dann gehen sie in die Gefahrenkarte, um zu sehen, was man gerechnet und was man betrachtet hat. Dort sind die Ansprüche auch gestiegen, weil die Gefahrenkarten auch für andere Sachen weiterverwendet werden. Das ist das eine. Dann merken sie, dass für einen bestimmten Bach, der bis jetzt pauschal betrachtet wurde, genauere Informationen nötig sind, weil sie es vielleicht weiterverwenden. Das ist das eine. Und das andere ist sicher auch fachlich. Wie zum Beispiel das RAMMS. Als das RAMMS rauskam, dann war klar: Wenn man neue Grundlagen hat. Oder auch bei Hangmuren hat man einmal die Grundlagen neu überdenkt. Sachen, welche bis dann gehinkt haben. Und dann muss man das genauer betrachten und es nach diesen Richtlinien machen.
JB	*Du hast die Weiterverwendung der Gefahrenkarte erwähnt. Was ist der Zweck oder die Verwendung der Gefahrenkarten?* *You mentioned the further usage of hazard maps. What is the purpose of hazard maps?*
E6	Sicher am meisten für die beiden Bereiche, welche ich schon erwähnt habe. Das eine sind die Wasserbauprojekte oder sonstige Bauprojekte, bei denen man nicht weiss, ob sie kommen, aber vielleicht wird einmal irgendwo eine Strasse geplant. Plötzlich merkt man, da muss man Steinschlagnetze installieren und da muss man wissen, welche Art von Felsen dort kommen können. Was haben diese für Dimensionen und welche Netze sind nötig? Und das andere im Wasserbau ist genau das gleiche. Das zweite ist noch für die Schutzmassnahmen an Objekten. Dann kann man sagen… Wenn man eine unscharfe Gefahrenkarte hat, weil man pauschal gerechnet hat, und man dann die Lage für ein Gebäude beurteilen muss, dann weiss man gar nicht genau, ob ein Ereignis von diesem oder jenem Bach kommt, weil sie zusammenfliessen. Dann braucht man eine Intensitätskarte, bei welcher man jede Prozessquelle einzeln einblenden kann. Somit weiss man eher, was man machen muss. Sonst gibt es immer Fragen und Rückfragen. Jemand kann eine Beurteilung von der anderen Seite des Baches machen und sagen: Die anderen Werte stimmen gar nicht. Wie kommt man auf so was? Und derjenige erinnert sich auch nicht mehr genau. Dann muss man das alles aufarbeiten. Das sind die Hintergründe.
JB	*Bezgl. Farbschema: Wenn das Farbschema der Gefahrenkarten nicht standardisiert wäre, würdest du ein anderes verwenden? Wenn ja, welches?* *Regarding the colour scheme: If the colour scheme of hazard maps was not standardized, would you use a different one? If yes, what kind?*
E6	Ehrlichgesagt finde ich es ziemlich gut. Ich habe es gerade vor 2 Stunden in der Pause gesagt. Wenn ich mit Wasserbau-Leuten zu tun hatte, als ich an der E. arbeitete, drehte ich fast durch, weil sie sich akribisch auf Kommastellen genaue Aussagen gewohnt sind. Man hat einen Niederschlag oder eine Abflussstatistik und dann weiss man haargenau, was ein 100-jährliches Ereignis ist, Nachkommastellen genau. Als ich an der Uni war, war ich bei H.K., er war mein Professor und eine Koryphäe in der Entwicklung der Naturgefahren(karte) auch in der Farbgebung zumindest meines Wissens. Und er hat uns immer mit auf den Weg mitgegeben: Es geht hier um Grössenordnungen. Man kann die Natur nie abschliessend erfassen. Und wenn man ein Szenario so definiert und es so genau beschreibt und es auf einer Karte so genau darstellen würde, was ist dann, wenn eine Verklausung unerwartet weiter oben auftritt, weil eine Wurzel anders liegt oder man

	einfach Pech hat? Und dann kommt es weiter, weil es nicht genug allgemein gefasst war. Klar soll man es nicht zu allgemein machen, man will auch eine Aussage machen und nicht alles unscharf machen. Aber irgendwo braucht es Platz, dass zwei Beurteiler auf ein ähnliches Resultat kommen. Es muss nachvollziehbar sein und nicht, dass die Karte anhand von Miniatur Anpassungen schon nicht mehr stimmt. Insbesondere auch, das wollte ich vorher bei der Verwendung durch den Kanton noch sagen, eine grössere Detaillierung. Wenn ich nämlich meine eigenen Gefahrenkarten nach 10 Jahren, nach einem Ereignis oder wegen einer geplanten Massnahme revidieren muss: Dann muss man nachvollziehen, was man dazumal gedacht hat, damit man darauf aufbauen kann und mit der Massnahme plausible Antworten hat, weshalb sie schützt. Und wenn jemand, der mal hier war, nicht mehr auf dem Beruf ist, dann ist es schwierig, wenn es zu genau wurde. Beispielsweise, wie du gesagt hast, wenn man eine zu genaue Farbgebung hat. Dann wird es schwierig zu wissen, was der Vorgänger sich überlegt hat. Das sind dann solche Miniatur Nuancen, dass es dann ein enormer Aufwand ist. Und ob es besser wird oder eher stimmt, ist die andere Frage. Darum denke ich, dass es nicht so schlecht ist.
JB	*Wieder eher zur Verwendung der Gefahrenkarten: Welche Entscheidungen können mithilfe der Gefahrenkarten getroffen werden?* *Coming back to the usage of hazard maps: What are decisions which can be made with the help of hazard maps?*
E6	Zum Beispiel in der Raumplanung? (JB: *Ja, genau.*) In der Raumplanung: darf man bauen oder darf man nicht bauen? Man sagt immer: Die Massnahmenplanung muss bei der Raumplanung anfangen. Man weicht lieber einfach aus. Und wenn das nicht geht, dann beginnt man über Schutzmassnahmen, bauliche Massnahmen zu reden. Und die zweite Verwendung habe ich auch schon angetönt, ist indirekt in der Gefahrenkarte für die Dimensionierung von Massnahmen aus dem Bericht zur Gefahrenkarte. Das ist im weitesten Sinn Raumplanung: die Linienführung von Strassen, dass man nicht mitten durch das Gefahrengebiet geht. Das gab es an gewissen Orten. Gerade auch in Guttannen in Richtung Wallis, also das ganze Tal über den Grimsel. Dort haben wir für Swissgrid, die Stromleitungen, eine Grobanalyse durchgeführt, um herauszufinden, auf welche Talseite sie am besten die Masten bauen. Da haben wir alle Gefahren behandelt, um zu sehen, wo die Wahrscheinlichkeit am kleinsten ist. Für solche Sachen werden die Gefahrenkarten gebraucht. Für Entscheidungen zur Planung von Leitungen und Strassen. (…) Wenn ich im S. eine hätte… Denn momentan ist im S. eine Baustelle. Nach diesem grossen Ereignis, durch das ich unseren Termin absagen musste, das war übrigens das grösste überhaupt. Ja, das war das grösste überhaupt bis jetzt. Dort haben sie in der Aare jetzt eine Baustelle, welche ich jeden Morgen freigebe. Und dort wäre ich froh, wenn es eine Gefahrenkarte gäbe, sodass ich sagen könnte, sie können arbeiten, sie sind ausserhalb des Gefahrenparameters. Ich denke gerade für die Planung bei Bauarbeiten, dass man überhaupt sieht, ob sie sich grundsätzlich in einem Gefahrengebiet befinden oder nicht. Dann wüsste man, ob man ein Warnsystem einrichten muss oder nicht.
JB	*Aber das ist jetzt auch eher das Problem, weil gerade ein Ereignis geschah, sodass sich die Situation geändert hat?* *So that is more of a problem because an event has just occurred which changed the situation?*
E6	Das stimmt, aber man hat es auch an anderen Orten. Beispielsweise ein Wasserkraftwerk, welches gebaut wird, dort besteht eine Lawinengefahr und sie arbeiten aber auch im Winter. Und dann muss man schauen, ob man am Morgen die Baustelle freigeben muss oder nicht. Und das sieht man auch in der Karte, ob sie überhaupt in der Ausbreitung sind. Welcher Teil der Baustelle ist okay, welchen muss man mit einem Alarmsystem schützen?
JB	*Dann zur Überleitung zu den Unsicherheiten: Was sind die grössten Unsicherheiten spezifisch bei Murgängen?* *Transitioning to the uncertainties: What are the biggest uncertainties in debris flows?*

E6	(…) Den Triftgletscher kennst du wahrscheinlich. Im Berner Oberland beim Sustenpass. Der hatte während gewisser Zeit eine Gletscherzunge über einem See, das war circa im Jahr 2005. Dort hat man befürchtet, dass ein Teil abbricht und einen Seeausbruch verursacht, was weiter unten Murgänge verursachen könnte. Das ist beispielsweise ein Fall, bei dem man sagte: Wie möchte man diesen Murgang einschätzen? Da ist man unsicher, ob etwas kommt, wie gross es kommt. Man kann es über das Verdrängungsvolumen der Abrisszone schätzen. Aber auch wie beim S. kann man die Jährlichkeit nicht festlegen. Es ist beides, ob er kommt – das ist die Jährlichkeit - und wenn er kommt: wie viel Material nimmt er mit? Und das kann man nicht in der Gefahrenkarte darstellen. Genau wie beim S., dass man die Jährlichkeit nicht angeben kann. Wenn es abbricht, ist es fertig. Und dort haben wir es so gemacht, dass wir über die Gefahrenkarte eine schraffierte Fläche gezeichnet haben. (**Bildschirm geteilt**) Das ist der Gletscher und dieser Teil konnte in den See abbrechen. Unten ist die normale Gefahrenkarte und das schraffierte ist die Überlagerung der Flutwelle. Das ist allerdings sehr Geschiebe-lastig, also schon fast Murgang-ähnlich. Das ist ein Beispiel, das mir in den Sinn kam bezüglich der Unsicherheit. Generell bei der Unsicherheit bei Murgängen ist es so, dass rein räumlich… Wenn es an mehreren Orten vorkommen kann, dann nimmt man die Umhüllung aller möglicher Szenarien, wo er ausbrechen könnte. Und dort ist die Unsicherheit, wo er ausbricht, und die Möglichkeit, dass er ausbrechen kann, drin. Und bei der Ausbreitung gibt es immer Unsicherheiten. Das eine ist… Wie sagt man… Also die Wahrheit ist kontinuierlich und die Karte hat dann eine kategorisierte Abstufung. Und dort muss man einen Cut machen, was nicht sehr einfach ist. Dann gibt es einen Strang, der weiter nach unten geht. Da muss man sagen, das generalisiert man. Es gibt zwar 35 mögliche Zungen, aber irgendwo muss man abschneiden. Das heisst ein einzelner Zug kann noch weiter gehen. Die Unsicherheit ist dort, ich könnte auch sagen, man kann noch eine Bandbreite definieren, bis wo sie (die Zungen) kommen. Aber ob diese dann stimmt, ist die andere Frage. Eigentlich kann man nicht oben die Elle messen und unten «milimetern», denn oben weiss man oft nur approximativ, wie viel Geschiebe kommt.
JB	*Dann macht man sich beim Prozess des Erstellens der Gefahrenkarte viele Gedanken zur Unsicherheit, aber man muss es dann abstrakt darstellen?* *So, during the process of creating the hazard maps a lot of thought is put into uncertainty, but one has to display it in an abstract way in the end?*
E6	Ja und die Kartographie ist das eine, aber im Bericht zu jedem Murgang-Graben wird vor allem beim Geschiebe bewusst eine Bandbreite integriert. Dass man sagt: im 30-jährlichen sind im Potenzial 1'000-2'000 Kubikmeter möglich, ob es kommt, ist die andere Frage. Also immer 2'00-4'000 oder 4'000-6'000. Dass man eine Bandbreite abbildet, denn es sind auch immer grosse Zeithorizonte. Und danach hat man eigentlich… Das war noch vor dem Modellieren. Bei der Modellierung ist es einfacher, da gibt man einen Wert ein und schaut, wie weit es kommt. Aber wenn man es nicht modelliert, nimmt man das Volumen und verteilt es in der Fläche. Man schaut, wo man in der Vergangenheit welche Ablagerungshöhen hatte, wie häufig diese Ereignisse stattfanden, ob es eher ein 30 oder ein 100-jährliches ist. Oder auch der Niederschlag, der dann jeweils fiel. Und dann sagt man: in diesem Gebiet gab es jeweils 0.5m. Wenn man so viel Material hat, wie weit nach unten muss es kommen, wenn man 0.5m Ablagerung hat? Da war es ein Herantasten und irgendwo zieht man einen Strich. Der Strich wird an einem Ort gezogen, der auch sonst plausibel scheint, beispielsweise bei einer Vertiefung oder einer Abflachung an einem steilen Hang. Dann wählt man solche Orte, weil man es nicht genauer darstellen kann. Aber sonst kann man es nicht auf der Karte darstellen. Und wir sagen auch, es ist parzellenscharf. Die Gefahrenkarte ist parzellenscharf. Und irgendwo in dieser Bandbreite bewegen wir uns.

JB	*Wenn man in die andere Richtung geht: Wenn man die Datenunsicherheit darstellen würde, wie es immer wieder diskutiert wird. Würde es den Entscheidungsträgern helfen oder würde es den Entscheidungsprozess komplexer machen?* *If one were to go in the other direction: If one would display data uncertainty, as it is frequently discussed. Would it help decision-makers, or would it make their decision-making process more complicated?*
E6	Ich glaube, es würde das Ganze sehr komplex machen. Es ist sicher eine Diskussion, welche ich mit dem Kanton führen würde. Zuerst intern und doppelt und dreifach nachprüfen, ob es heikel ist. Und wenn es sicher ist, dass es eine Gratwanderung ist, würde ich es mit dem Kanton diskutieren. Ich würde es aber nicht offiziell nach aussen kommunizieren. Ich bin sehr für Transparenz, aber im Entscheidungsprozess wird man sofort politisch. Sie kommen immer mit politischen oder persönlichen Anliegen. Und schlussendlich darf es keine politisch gefärbte Karte sein. Um dort die Neutralität nicht zu gefährden, würde ich die Leute dort aussen vor lassen. Ausser wenn es eine neutrale Frage ist, ob man in einem Bereich schon ein Ereignis beobachtet hat. Aber sonst würde ich diese Diskussion nur intern führen. Wir sind in der Regel an Orten… Wenn ich weiss, dass irgendwo etwas geplant wird, dann kommt es schon vor, dass man im Feld sagt, dass es blau ist, aber haarscharf an der Grenze. Dass man die Grenze bewusst so setzt, damit die Sensibilisierung gross ist. Dann ist es optisch auch an der Grenze. Zum Beispiel bei Lawinen, wenn man bei 30 kPa ist, was eine Grenze ist, bei der Staubeton nötig ist, ob man dann ein paar Meter davor oder dahinter ist, braucht es ihn. Sonst ist das Gebäude nicht geschützt. Dass man dort auf diese Weise kommuniziert, aber mehr, dass sie wissen, weshalb sie starke Bauten brauchen.
JB	*Dann geschieht dies eher im Prozess oder im Beratungsverfahren?* *So, this happens during the process or a consultation?*
E6	Ja, genau. Bei der Gefahrenkarte generell vermeidet man das Gespräch, wenn sie gezeichnet ist. Man erklärt das eher bei allem, was vorher diskutiert wurde, wo sie auch mitdiskutieren konnten. Beispielsweise weshalb eine Linie genau hier liegt. Aber wenn es um eine persönliche Beratung geht bei einem Gebäudebau, wenn man noch genau für ein Gebäude modelliert, dann findet diese Diskussion schon statt. Und den Leuten ist es auch klar, weil sie wollen auch nicht, dass ihr Haus kaputt geht. Meistens im Gegenteil. Die meisten sagen: Oh ja, dazumal kam es in einem Winter sehr nahe. Also sind sie sehr einsichtig.
JB	*Mit meinen vorbereiteten Fragen wäre ich dann durch. Gibt es noch einen Aspekt bezüglich der Gefahrenkarten, deiner Arbeit oder zu Murgängen, was du noch erläutern möchtest?* *Those were all the questions I prepared. Are there other inputs on hazard maps, your work and regarding debris flows that you would like to mention?*
E6	Ich frage mich gerade bezüglich der Klimaszenarien. Dort gibt es noch eine Unsicherheit, weil man weiss, dass das 20. Jhd. sehr ereignisarm war. Und das Jahrhundert vorher und das jetzt ist ereignisreicher. Statistisch blickt man immer auf eine Zeit zurück, in der man nicht wirklich vergleichbare Werte hat. Das muss man sich bewusst sein, dass man quasi einen Zuschlag geben muss. Zum Beispiel beim Tal Fluss Kander hiess es, dass man 10% auf die Abflüsse schlug aufgrund der Klimaunsicherheit oder im Wissen, dass es mehr Starkniederschläge geben wird. Schlussendlich ist das eine Diskussion, welche auch in die Unsicherheiten geht, aber man ist auf der sicheren Seite. Man diskutiert das in der frühen Phase der Szenarien-Bildung. Aber man stellt es in der Gefahrenkarte nicht im Sinn einer Bandbreite dar.
JB	*Sodass auch die Unsicherheiten bezüglich des Klimawandels ganz allgemein auch integriert werden speziell bei der Überarbeitung der Gefahrenkarten? Oder ist das eine zu grosse Herausforderung?* *So, the uncertainties relating to climate change in general are also part o fit especially during the revision of the hazard maps? Or is that a challenge that exceeds the capacity?*

E6	Man hat zwei Sachen. Man hat in der Gefahrenhinweiskarte, welche inklusive Klimawandel ist, drei Zeithorizonte integriert. Das waren 2030, 2060, 2100. Und man sagte, dass bis dann der Permafrost x viel höher ist, dann hat man y mehr Geschiebe, welches mobilisiert werden kann, und man hat solche Niederschläge mit dieser Temperaturzunahme bis dann. Danach hat man drei Murgang Modellierungen aus einem Gebiet, welche immer weiter runterkommen. Dort macht es Sinn. Aber bei der Gefahrenkarte ist es so, dass sie ca. alle 10 Jahre sowieso revidiert werden. Dann muss man es auf dem Radar haben, dass man allfällige neue Geschiebeherde berücksichtigt und beim Wasser etwas an die obere Grenze geht, wenn man eine Bandbreite hat. Daher macht es dort keinen Sinn, das 3-stufig abzubilden zum Beispiel. Noch etwas anderes. (**Bildschirm geteilt & Erklärung einer Präsentationsfolie**) Da sieht man für ein Einzugsgebiet verschiedene Abflussabschätzungen für Hochwasser. Wenn es ein 100-jährliches Ereignis ist, dann hat man verschiedene Methoden, um den Abfluss abzuschätzen, und dann noch die Maximalabflüsse zum Plausibilisieren. Mit einem kleinen Maximalabfluss wäre es komisch, wenn man einen höheren 100-jährlichen Abfluss hat. In diesem Beispiel ist es jetzt so heterogen, dass es kaum mehr brauchbar ist. Aber normalerweise hat man normalere Resultate und dann hat man auch dort eine Bandbreite. Dann sagt man, mit dem Klimawandel nehmen wir innerhalb der Bandbreite eher einen oberen Wert, anstatt einfach den Mittelwert zu nehmen. Auch auf diese Weise wir die Unsicherheit im Zusammenhang mit dem Klimawandel mit einbezogen. Das wollte ich noch sagen. Und auch das mit der Vergangenheit, dass es früher nicht so viele Ereignisse gab. Und auch beim Geschiebe, entweder man hat mehr Geschiebe oder weniger und mehr Abfluss oder weniger und dann ist das 100-jährliche in einer Grössenordnung von-bis. Genau weiss man es nicht.

G. RAMMS Debris Flow Modelling Workflow

The following screenshots provide more detailed information on the debris flow modelling using RAMMS:DEBRIS FLOW.

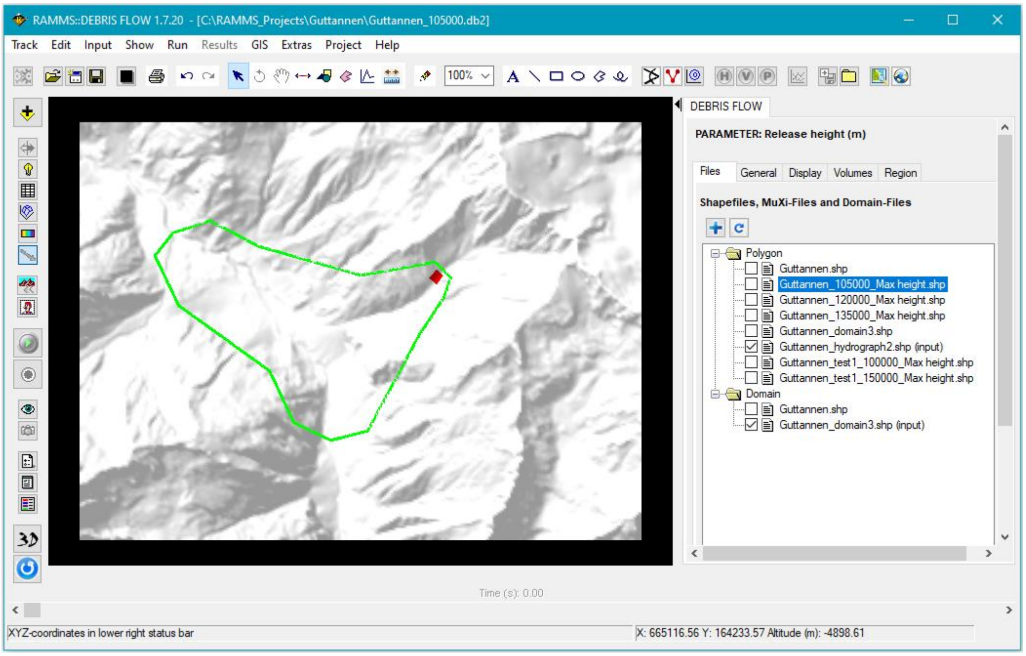


Figure 59: RAMMS::DEBRIS FLOW interface at location GUT showing the manually drawn domain (green) and the hydrograph (red).

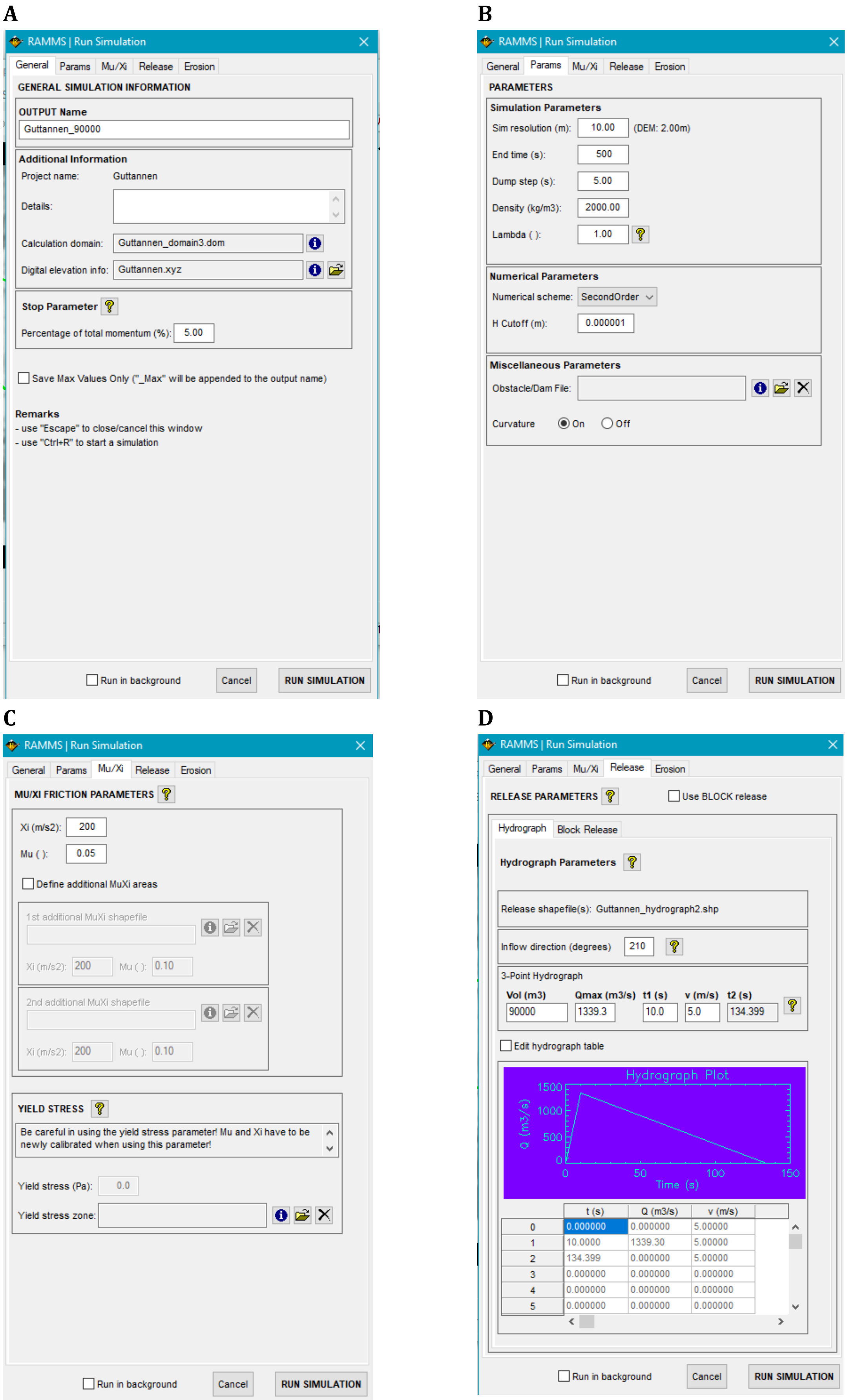


Figure 60: RAMMS::DEBRIS FLOW simulation run: Definition of an output name, the calculation domain and the DEM (A). Definition of simulation parameters including the resolution (B). Entering of the friction parameters (C). Definition of release parameters such as the debris flow volume and the direction (D).

H. QGIS Uncertainty Modelling Workflow

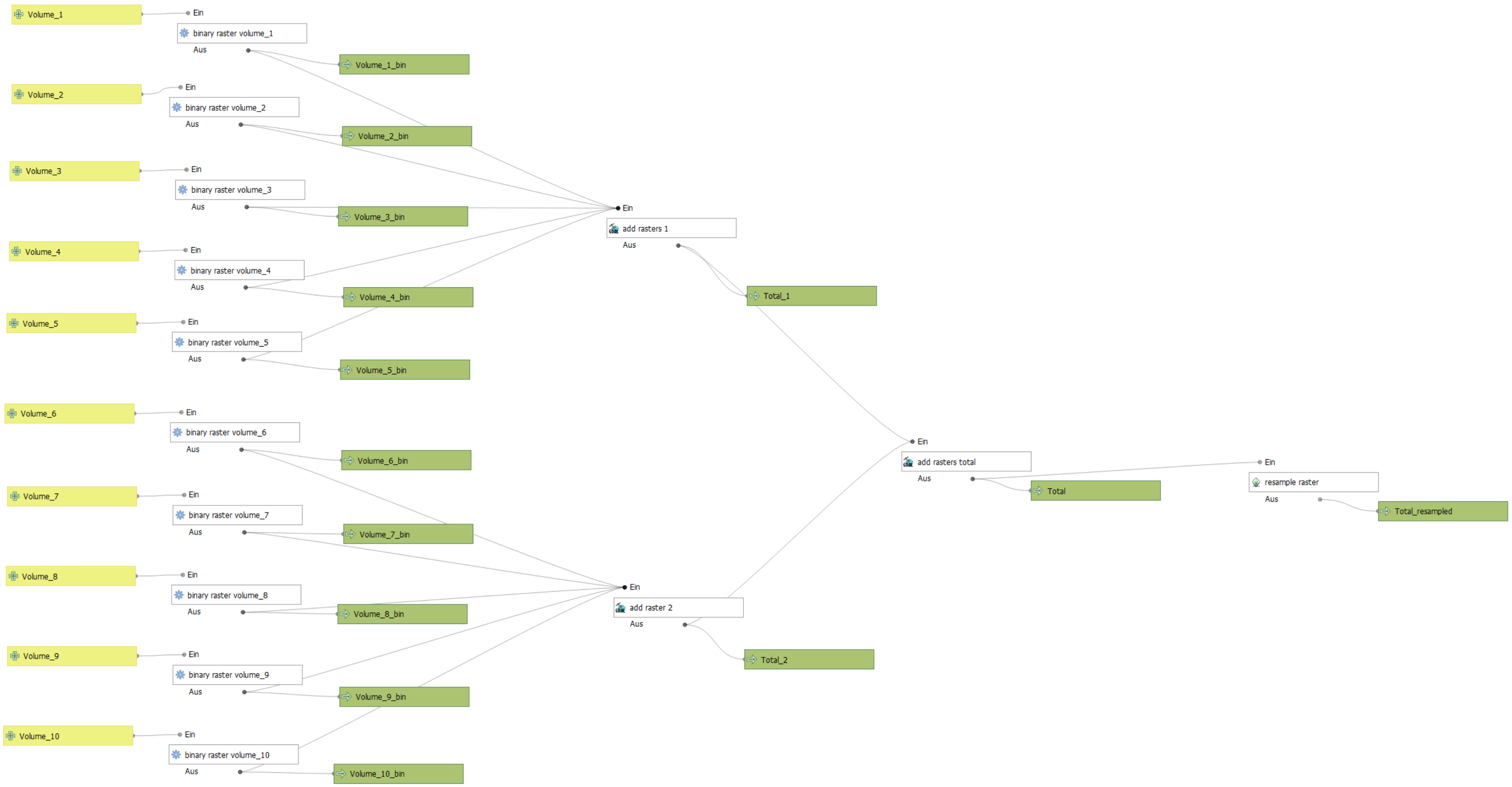


Figure 61: Model shows the GIS workflow performed in QGIS: Debris flow volume rasters were transformed into binary masks. The rasters were added in two groups, which were then summed up. Lastly, the raster was resampled to a resolution of 5m to create a smoothed uncertainty dataset.

I. Trial Planning

The following section provides details on the planning of the house locations in the map stimuli:

An analysis of the slopes at all debris flow locations showed that slopes ranged from 0° to 89° with an overall mean slope of 17.8° (slope values were extracted from the debris flow simulation with the largest spatial extent per location). Therefore, slope values were parted into the three categories: low (0 to 15°), medium (15 to 30°) and high (30 to 90°) (see Table 28).

Table 28: Slope categories for the tested house locations including the number of trials each participant performed per slope category.

Slope category	**Slope value range**	**Number of trials per person**
Gentle	0 to 15°	6
Moderate	15 to 30°	8
Steep	30 to 90°	6

As the size of the debris flow sites can be quite different, the distance categories were not based on actual values, but were defined relative to each site (see Table 29). High distance locations were planned so that half of them represent locations right inside the debris flow uncertainty shape and half of them are right outside the shape. This was later used to investigate a potential boundary effect.

Table 29: Distance categories for the tested house locations including the number of trials each participant performed per distance category.

Distance category	**Distance definition**	**Number of trials per person**
Short	Close to riverbed	6
Medium	Half way between riverbed & boundary	8
Long	Either right inside or right outside the extent boundary	6 (3 inside, 3 outside)

Based on these two categorizations, all possible combinations of slope and distance were defined (see Table 30). These were then distributed as evenly as possible into 20 locations across the ten debris flow sites (see Table 31).

Table 30: Combinations of slope and distance categories including the number of trials per combination.

Distance / **Slope**	Short	Medium	Long
Gentle	2	2	2
Moderate	2	4	2
Steep	2	2	2

Table 31: Trial planning of house locations for each debris flow site including the combination of slope and distance category.

Debris Flow Site	**House Location Nr.**	**Slope Category**	**Distance Category**
Blenio (BLE)	1	Gentle	Short
	2	Steep	Medium
Bondo (BON)	3	Steep	Medium
	4	Gentle	Long
Brienz (BRZ)	5	Moderate	Short
	6	Steep	Long
Guttannen (GUT)	7	Moderate	Medium
	8	Steep	Long
Guttannen 2 (GUT2)	9	Steep	Short
	10	Moderate	Long
Lauterbrunnen (LTB)	11	Gentle	Medium
	12	Moderate	Short
Leuk (LUK)	13	Moderate	Medium
	14	Gentle	Long
Lütschental (LUT)	15	Steep	Short
	16	Moderate	Medium
Randa (RND)	17	Gentle	Short
	18	Moderate	Medium
Silenen (SIL)	19	Gentle	Medium
	20	Moderate	Long

J. Tobii Pro Lab: Experiment Design

The following figures give an insight into the exact configurations defined in Tobii Pro Lab.

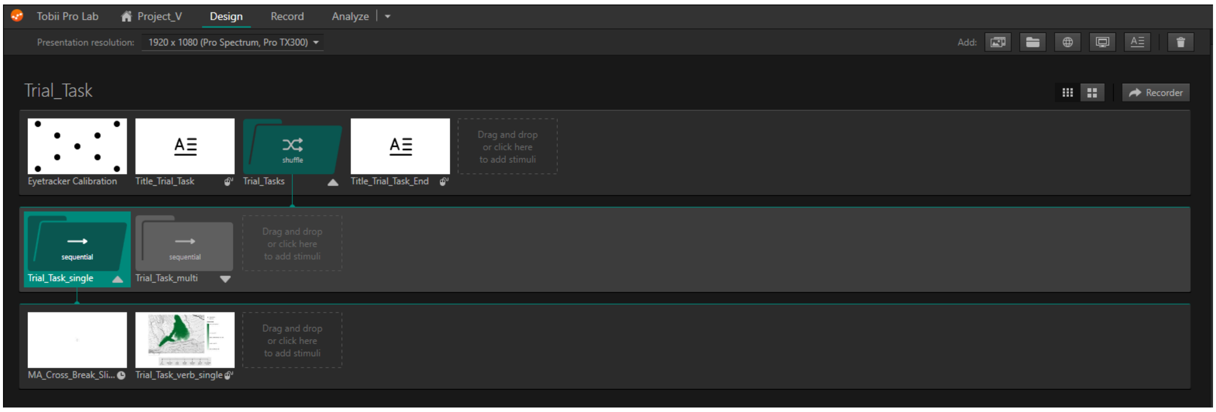


Figure 62: Design of the trial task timeline in Tobii Pro Lab for the verbal group.

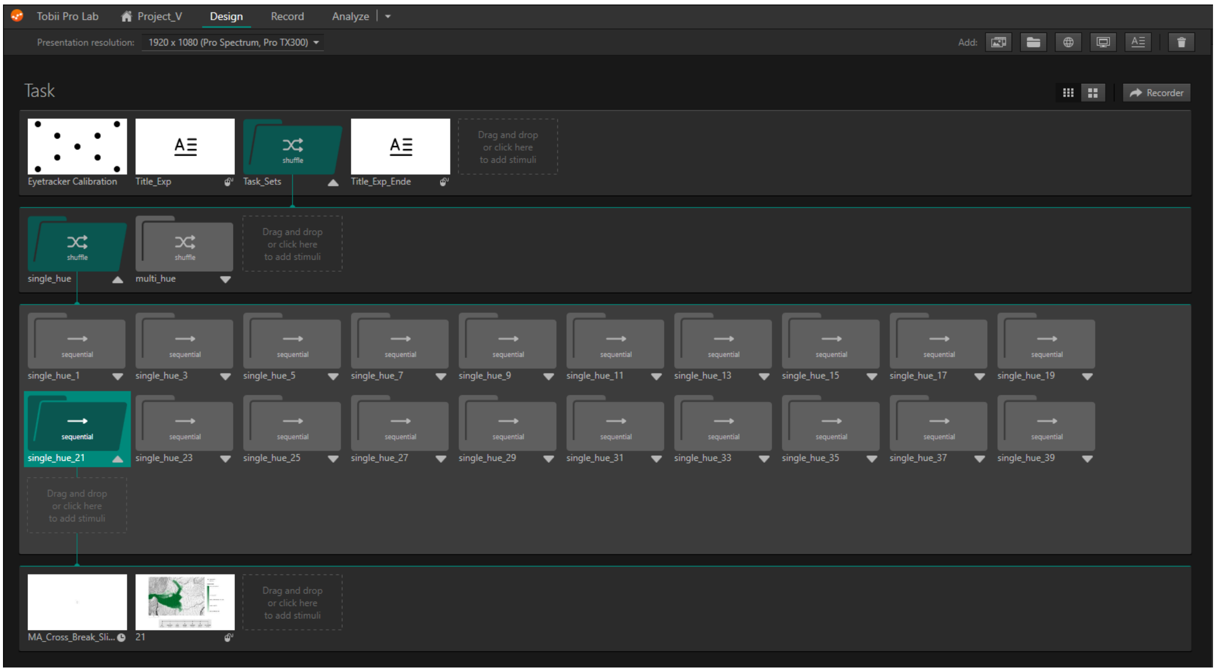


Figure 63: Design of the task timeline in Tobii Pro Lab for the verbal group.

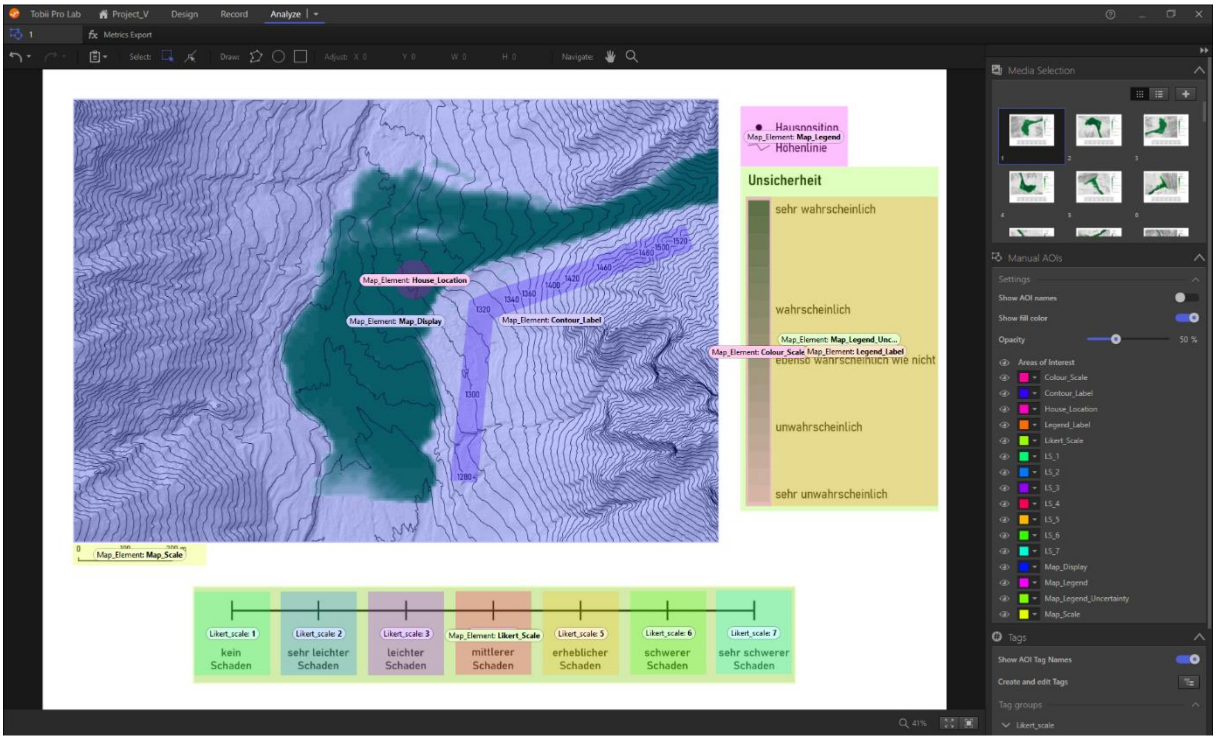


Figure 64: AOI tool where AOIs were defined for each map stimuli.

K. Recruiting

The following recruiting message was sent to potential participants via email:

Subject: Einladung zur Teilnahme an Studie für Masterarbeit – Jana Bracher

Liebe/r XY

Wie du vielleicht weisst, arbeite ich momentan an meiner Masterarbeit am Geographischen Institut der Universität Zürich in der Forschungsgruppe GIVA (Geographic Information Visualization and Analysis). In meiner Masterarbeit untersuche ich die Darstellung und Kommunikation von Unsicherheiten in der Vorhersage von Murgang-Ereignissen.

Gerne möchte ich dich als Studienteilnehmer/in einladen! Es handelt sich um eine Computer-basierte Studie im Eye Movement Lab (Y25-L-09) auf dem Campus Irchel der Universität Zürich (Winterthurerstrasse 190, 8057 Zürich). Die Studie dauert ca. 1 Stunde. Deine Aufgabe wird es sein, kurze Entscheidungsaufgaben aufgrund von Kartendarstellungen zu lösen, währenddessen deine Augenbewegungen mittels Eye Tracking Technologie aufgezeichnet werden.

Unter folgendem Link findest du eine Auswahl an freien Terminen für deine Teilnahme:

Doodle link

Bitte trage dich im Doodle mit Vor- & Nachname sowie deiner Emailadresse ein. Deine Informationen werden selbstverständlich vertraulich behandelt. Sollte dir keiner der freien Termine passen, darfst du dich gerne direkt bei mir melden, sodass wir gemeinsam einen anderen Termin finden.

WICHTIG: Die Durchführung des Studie ist der Hauptbestandteil meiner Masterarbeit, daher bitte ich dich, deine Terminzusage als definitiv zu sehen.

Voraussetzungen zur Teilnahme: gültiges COVID-Zertifikat, normale oder korrigierte Sehkraft (Kontaktlinsen/Brille), mind. 18 Jahre.

COVID-Schutzkonzept: 2G-Regel, Maskenpflicht, regelmässiges Lüften & Oberflächendesinfektion.

Es freut mich, wenn du an meiner Studie teilnimmst und so einen extrem wichtigen Beitrag zu meiner Masterarbeit beiträgst. Vielen Dank für deine Hilfe im Voraus!

Natürlich darfst du den Doodle-Link und alle Informationen auch mit weiteren Interessierten teilen. 😊

Liebe Grüsse

Jana Bracher

L. Consent Form

Teilnehmerinformation & Einwilligungsformular

Masterarbeit – Visualisierung & Kommunikation von Unsicherheiten bei Murgängen

Januar 2022

Teilnehmer-Nummer: ____

Zweck der Studie

Du bist eingeladen, an einer Studie zur «Visualisierung & Kommunikation von Unsicherheiten bei Murgängen» teilzunehmen. Wir möchten mit dieser Studie neue Erkenntnisse gewinnen über die Darstellung von Datenunsicherheit und deren Auswirkung auf räumliche Entscheidungen bei Murgang-Karten. Dieses Experiment wird von Jana Bracher im Rahmen ihrer Masterarbeit durchgeführt. Die Masterarbeit wird von Prof. Dr. Sara Fabrikant und Dr. Ian-Tanner Ruginski am Geographischen Institut der Universität Zürich geleitet.

Ablauf der Studie & damit verbundene, mögliche Risiken

Falls du dich entscheidest, an der Studie teilzunehmen, wirst du zuerst gebeten, einen Fragebogen mit Angaben zu deiner Person auszufüllen, gefolgt von einem kurzen theoretischen Einführungsteil. Weiter wirst du eine Serie von Aufgaben mit Karten am Computer bearbeiten. Während dieses Frageblocks mit den Karten werden deine Augenbewegungen aufgezeichnet. Das ist für dich völlig schmerzfrei und ungefährlich. Nach diesem Frageblock wirst du gebeten, einen weiteren Fragebogen auszufüllen. Folgende Daten werden anonymisiert aufgenommen: Antworten zu den gestellten Fragen in den Fragebögen und dem Aufgabenteil mit den Karten, die Antwortzeiten und deine Augenbewegungen. Der Versuch findet im Raum 25-L-9 («Eye Movement Lab») des Geographischen Instituts an der Universität Zürich statt. Er dauert ca. 60 Minuten und beinhaltet keinerlei Risiken für dich.

Vertraulichkeit der Daten

Alle Informationen, die während der Studie mit dir in Verbindung gebracht werden können, werden vertraulich behandelt und nur mit deiner ausdrücklichen Erlaubnis an Dritte weitergegeben. Mit deiner Unterschrift erlaubst du uns, die anonymisierten Ergebnisse des Versuchs mehrfach zu publizieren. Dabei werden keinerlei Informationen veröffentlicht, die es ermöglichen, dich zu identifizieren.

Bekanntgabe der Ergebnisse

Wenn du über die Ergebnisse der Studie auf dem Laufenden gehalten werden möchtest, bitten wir dich, dem Versuchsleiter oder der Versuchsleiterin deine Anschrift zu hinterlassen. Eine Kopie von zukünftigen Publikation(en) wird dir daraufhin zugestellt.

Einwilligung

Deine Entscheidung, an der Studie teilzunehmen oder nicht, wird zukünftige Beziehungen mit der Universität Zürich nicht beeinträchtigen. Entscheidest du dich dafür, an der Studie teilzunehmen, steht es dir jederzeit frei, die Teilnahme ohne Begründung abzubrechen. Solltest du Fragen haben, zögere bitte nicht, uns diese zu stellen. Sollten zu einem späteren Zeitpunkt Fragen aufkommen, wird Jana Bracher (jana.bracher@uzh.ch) oder Prof. Dr. Sara I. Fabrikant (044 635 51 50, sara.fabrikant@geo.uzh.ch), diese gerne beantworten. Du erhältst eine Kopie dieses Dokuments.

Mit deiner Unterschrift bestätigst du, obenstehende Informationen gelesen und verstanden zu haben und willigst ein, unter den dort beschriebenen Bedingungen am Experiment teilzunehmen.

Unterschrift des/der Teilnehmenden:	Unterschrift der Experimentleiterin:
___________________________________	___________________________________
Vor- & Nachname in Blockschrift:	Vor- & Nachname in Blockschrift:
___________________________________	___________________________________
Ort, Datum:	
___________________________________	

Widerruf der Einwilligung

Hiermit möchte ich meine Einwilligung, an der oben beschriebenen Studie teilzunehmen, widerrufen.

Unterschrift des/der Teilnehmenden:	Vor- & Nachname in Blockschrift:
___________________________________	___________________________________
Ort, Datum:	
___________________________________	

Mit dem Widerruf der Einwilligung beeinträchtigst du in keiner Weise deine Beziehung mit der Universität Zürich. Der Widerruf kann jederzeit und ohne Angabe von Gründen beantragt werden.

Den Widerruf der Einwilligung bitte an Prof Dr. Sara I. Fabrikant, Geographische Informationsvisualisierung und Analyse, Geographisches Institut, Universität Zürich, Winterthurerstr. 190, 8057 Zürich senden.

M. Pre-Test Questionnaire

Master Arbeit: Visualisierung und Kommunikation von Unsicherheiten bei Murgängen - Fragebogen 1 (N)

Zwischengespeicherte Umfrage laden

Master Arbeit: Visualisierung und Kommunikation von Unsicherheiten bei Murgängen - Fragebogen 1 (N)

Vielen Dank, dass Du an meiner Studie teilnimmst! Deine Teilnahme ist für meine Masterarbeit von zentraler Bedeutung und hilft, weitere Erkenntnisse im Bereich der Unsicherheitsvisualisierung zu sammeln.

Die Studie ist in 4 Teile aufgeteilt:

1. Fragebogen zu Deiner Person & Deinem Vorwissen
2. Thematische Einführung
3. Aufgabenblock mit Karten
4. Fragebogen zu Deinen Erfahrungen beim Lösen der Kartenaufgaben

Du wirst alle Teile der Studie an diesem Bildschirm absolvieren. Für den Aufgabenblock mit den Karten wird eine Eye-Tracking-Technologie verwendet, welche aber keinerlei Risiken für Dich birgt. Falls Du im Verlauf der Studie fragen hast, darfst Du Dich jederzeit melden.

Nun werden Dir einige Fragen zu Deiner Person und Deinem Wissensstand gestellt. Darauf folgt eine kurze thematische Einführung in den Prozess des Murgangs.

Weiter

Teil 1

* Diese Frage wird von der Experimentleiterin ausgefüllt!

Teilnehmer/innen-ID:

Deine Antwort muss mindestens 0 sein.

In diesem Feld kann nur ein ganzzahliger Wert eingetragen werden.

Teil 2

* Wie alt bist du?

Deine Antwort muss mindestens 0 sein.

In diesem Feld kann nur ein ganzzahliger Wert eingetragen werden.

*** Was ist dein Geschlecht?**

Bitte wähle eine der folgenden Antworten:

Falls Du 'Sonstiges:' auswählen solltest, spezifiziere bitte Deine Auswahl im entsprechenden Textfeld.

- ○ Weiblich
- ○ Männlich
- ○ Sonstiges: ______

*** Was ist deine höchste abgeschlossene Ausbildung?**

Bitte wähle eine der folgenden Antworten:

Falls Du 'Sonstiges:' auswählen solltest, spezifiziere bitte Deine Auswahl im entsprechenden Textfeld.

- ○ Kein Abschluss
- ○ Obligatorische Schule
- ○ Sekundär - Berufsausbildung EFZ, Maturität
- ○ Tertiär - Höhere Berufsbildung (Höhere Fachschule, Höhere Fachprüfung)
- ○ Tertiär - Hochschule (Fachhochschule, Universitäre Hochschule)
- ○ Sonstiges: ______

*** Arbeitest du in einem dieser Fachbereiche: Naturgefahren, Gefahrenkarten, Kartografie, GIS (Geographische Informationssysteme)? Wenn ja, schreibe deine Berufsbezeichnung in das Kommentarfenster.**

Bitte wähle eine der folgenden Antworten:

- ○ Ja
- ○ Nein

Bitte gib hier Deinen Kommentar ein: ______

*** Wurde bei dir durch eine Fachperson (Optiker, Augenarzt) je eine Sehschwäche festgestellt?**

Bitte wähle die zutreffenden Punkte aus.

- ☐ Ja, Farbenblindheit
- ☐ Ja, Brille oder Kontaktlinsen
- ☐ Nein
- ☐ Sonstige Sehschwächen: ______

Teil 3

*** Wie oft (in der Freizeit oder bei der Arbeit) verwendest du Karten (mobile oder gedruckte Karten)?**

Bitte wähle eine der folgenden Antworten:

Falls Du 'Sonstiges:' auswählen solltest, spezifiziere bitte Deine Auswahl im entsprechenden Textfeld.

- ○ Täglich
- ○ Wöchentlich
- ○ Monatlich
- ○ Jährlich
- ○ Sonstiges: [ ]

*** Hast du bereits Erfahrung mit dem Erstellen von Karten?**

Dies bezieht sich auf die Erstellung einer Karte im Rahmen eines Kartografie-/GIS-Kurses oder privat.

✔ Ja | ⊘ Nein

*** Wie gross ist deine Erfahrung in folgenden Bereichen?**

Ein Wert von 1 bedeutet, dass du noch wenig Erfahrung in diesem Bereich hast. Die 5 heisst, dass du bereits grosse Erfahrung in diesem Bereich hast.

	1	2	3	4	5
Karten lesen	○	○	○	○	○
Kartographie	○	○	○	○	○
Geographische Informationssysteme (GIS)	○	○	○	○	○
allgemeiner Umgang mit Datenunsicherheit	○	○	○	○	○
Statistik	○	○	○	○	○
Gefahrenkarten	○	○	○	○	○
Murgänge	○	○	○	○	○

*** Warst du bereits persönlich von einem Murgang betroffen (Schäden, Einschränkungen in deiner Mobilität etc.)? Wenn ja, erläutere deine Erfahrung kurz im Kommentarfeld.**

Bitte wähle eine der folgenden Antworten:

- ○ Ja
- ○ Nein

Bitte gib hier Deinen Kommentar ein: [ ]

N. Thematic Introduction

Thematische Einführung

Was ist ein Murgang?:

Ein Murgang ist ein natürliches Hangrutsch-Phänomen, bei dem sich ein Gemisch aus Wasser und Sediment den Hang hinunter bewegt. Da Murgänge eine Mischung aus Überschwemmungen und Massenbewegungen sind, wird der Murgang in die Naturgefahrenkategorie der Wasserprozesse eingeteilt. Die Murgangsaison dauert in der Schweiz in der Regel von Mai bis Oktober.

Abbildung 1: Heranschreitender Murgang. (Bild: Corina Gwerder WSL/ETH).

Faktoren:

Es gibt drei Voraussetzungen, die in Kombination einen Murgang auslösen können:

1. Steiles Terrain
2. Lockeres Sediment
3. Wasser (von einem Starkregenereignis oder einem Gletscherseeausbruch)

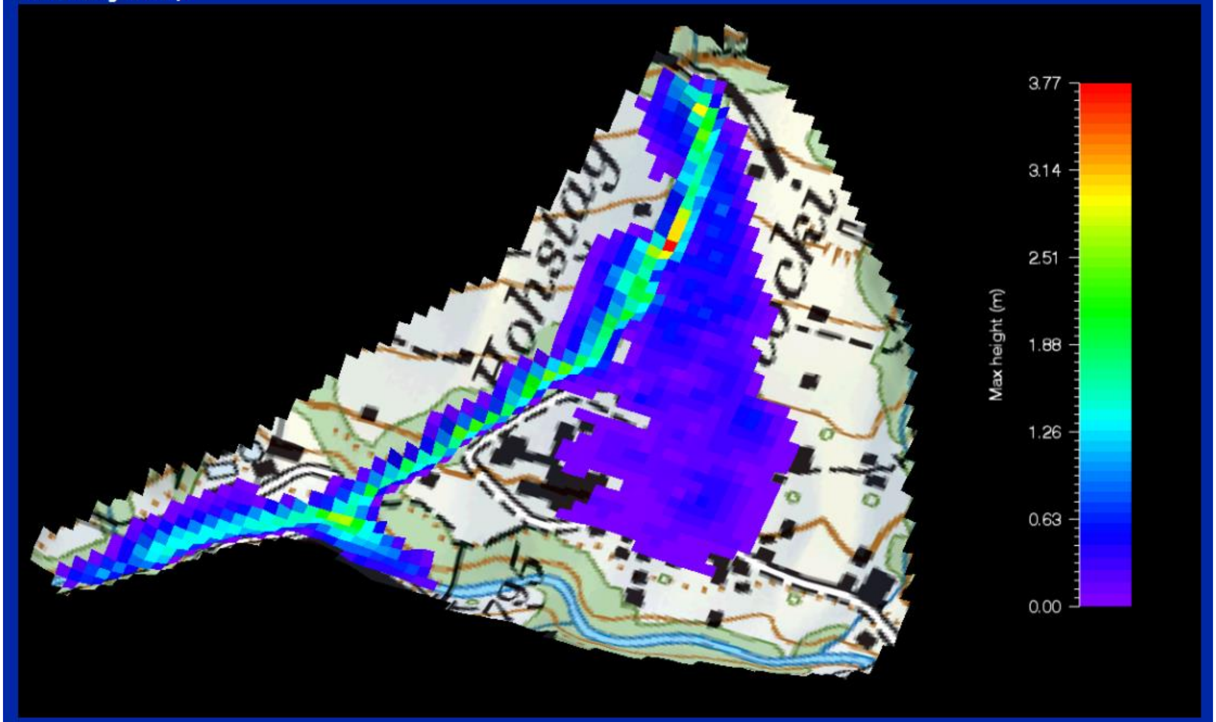
Gefahrenpotenzial von Murgängen:

Aufgrund ihrer destruktiven Kräfte können Murgänge grosse Schäden an der Siedlungs- oder Verkehrsinfrastruktur bewirken. Durch ihre hohe Geschwindigkeit von bis zu 40-60 km/h stellen Murgänge auch für den Menschen selbst eine grosse Gefahr dar. Der Umgang mit dieser Gefahr ist ein zentrales Element des Naturgefahren-Managements.

Unsicherheiten in der Murgang-Vorhersage:

Ehemalige Murgang-Ereignisse und Modell-Simulationen (siehe Abbildung 2) können helfen, die Gebiete, die von zukünftigen Murgängen betroffen sein könnten, zu erkennen. Trotzdem ist es schwierig, die räumliche Ausbreitung eines Ereignisses präzise vorherzusagen. Denn es gibt viele Unsicherheiten in den vorhandenen Daten und den Modellrechnungen von Murgängen, speziell bezüglich des potenziellen Volumens eines Murgangs. Zusätzliche Unsicherheiten treten aufgrund des fortschreitenden Klimawandels auf. Durch die ansteigende Lufttemperatur und die Veränderung des Niederschlagsmusters in der Schweiz, könnte es zu einer Veränderung der Häufigkeit von Murgängen kommen. Zudem ist es möglich, dass vermehrte Extrem-Niederschlagsereignisse die Grösse von Murgang-Ereignissen erhöhen.

Abbildung 2: Modellsimulation eines Murgang-Ereignisses. Die Farblegende beschreibt die maximale Höhe der Sedimentablagerung (auch Sedimentation genannt).



Falls du hierzu noch Fragen hast, darfst du dich gerne bei der Experimentleiterin melden.

O. Introduction to the Task

Einführung in die Aufgabenstellung

Szenario:

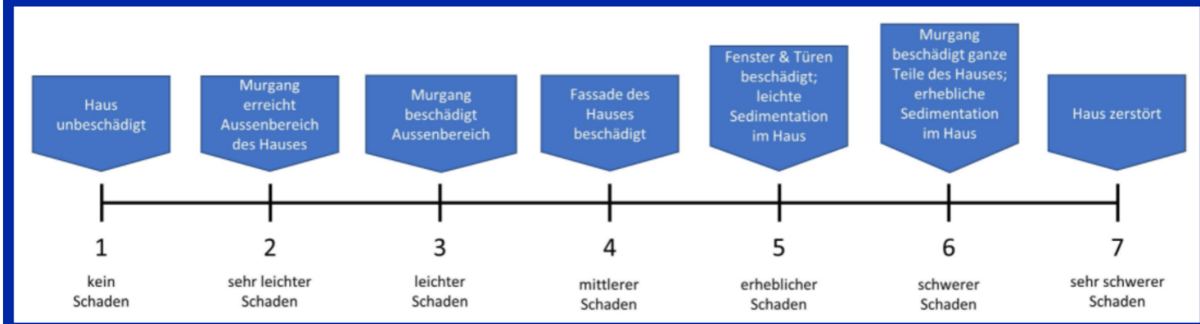
In Deiner Rolle als Mitglied der Gemeindeverwaltung bist Du für das Management der Naturgefahren im Gemeindegebiet verantwortlich. Murgangereignisse haben, wie bereits erwähnt, ein hohes Schadenpotenzial. In Deiner Gemeinde gibt es eine Vielzahl an potentiellen Murgang-Gefahrenzonen.
Für die ausgeschiedenen Gefahrenzonen, in welchen Murgänge auftreten könnten, wurden Modellierungen durchgeführt, um die räumliche Ausbreitung eines Murgangs vorherzusagen. Diese Vorhersage ist jedoch mit Unsicherheiten verbunden.

Aufgabe:

In einer Serie von Karten werden Dir die Resultate der Modellrechnungen, inklusive der berechneten Unsicherheiten, von zukünftigen Murgängen und deren Ausbreitungen gezeigt. Auf diesen Karten ist jeweils auch die Lage eines betroffenen Hauses mit einem Punktsymbol dargestellt. Deine Aufgabe ist es nun, den Schaden am betroffenen Haus auf einer Skala von 1 (kein Schaden) bis 7 (Haus zerstört) einzuschätzen.

Schaden-Skala:

Damit Du Dir die Bandbreite der möglichen Schäden, die Murgänge anrichten können, vorstellen kannst, findest Du unten ein Beispiel der Skala. Diese Skala wird Dir zur Lösung der Karten-Aufgaben zur Verfügung stehen. Sedimentation bedeutet, dass Material, welches vom Murgang mitgetragen wird, im Haus abgelagert wurde.



Falls du hierzu noch Fragen hast, darfst du dich gerne bei der Experimentleiterin melden.

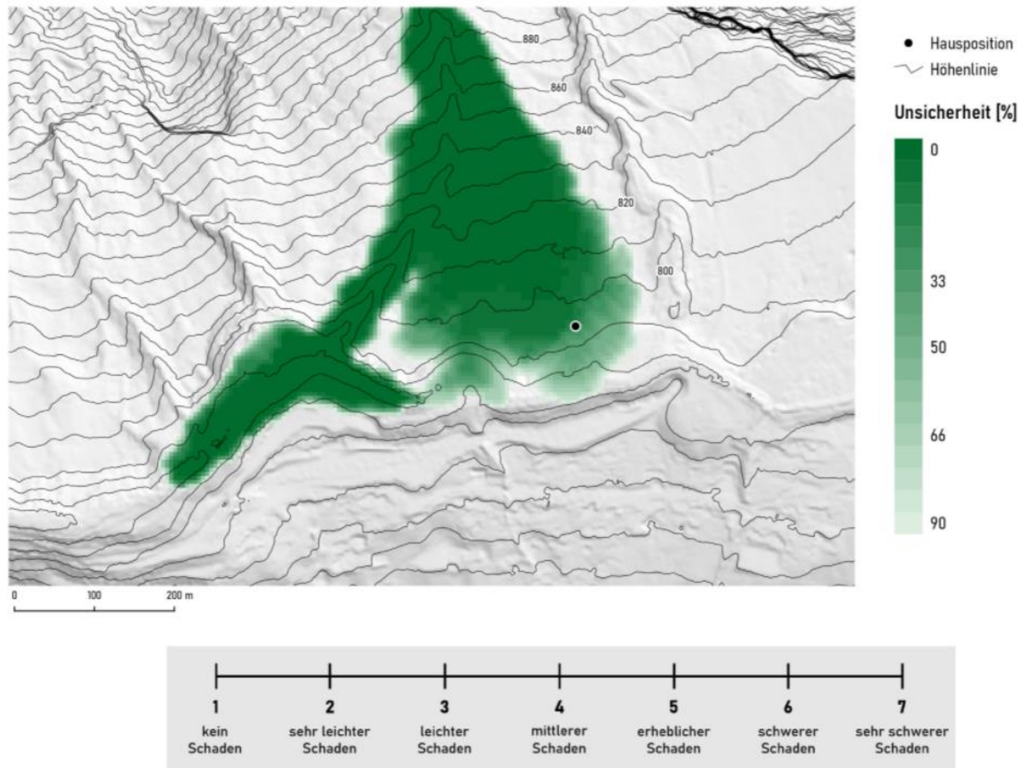
Einführung in die Aufgabenstellung

Die Abbildung unten zeigt ein Kartenbeispiel eines Modellierungsresultats der Murgangsausbreitung inklusive deren Unsicherheit, wie Du sie während des Aufgabenblocks sehen wirst. Die Lage eines betroffenen Hauses ist mit einem Punktsymbol in der Karte gekennzeichnet. Die Basiskarte zeigt das betroffene Gelände. Die Symbolisierung des Karteninhalts ist in der Legende rechts erklärt.

Die Höhenlinien dienen Dir zur Interpretation des Geländes. Beachte jeweils die Beschriftungen der Höhenlinien, denn ein Murgang fliesst stets von einer hohen zu einer tiefen Lage. Liegen die Höhenlinien nahe zusammen, bedeutet dies, dass das Gelände steil ist. Wenn grössere Abstände zwischen den Höhenlinien liegen, ist das Terrain flach.

In den Karten des nachfolgenden Aufgabenblocks bitte ich Dich jeweils, den zu erwartenden Schaden am betroffenen Haus in der Karte mithilfe der bereits besprochenen Schadensskala einzuschätzen.

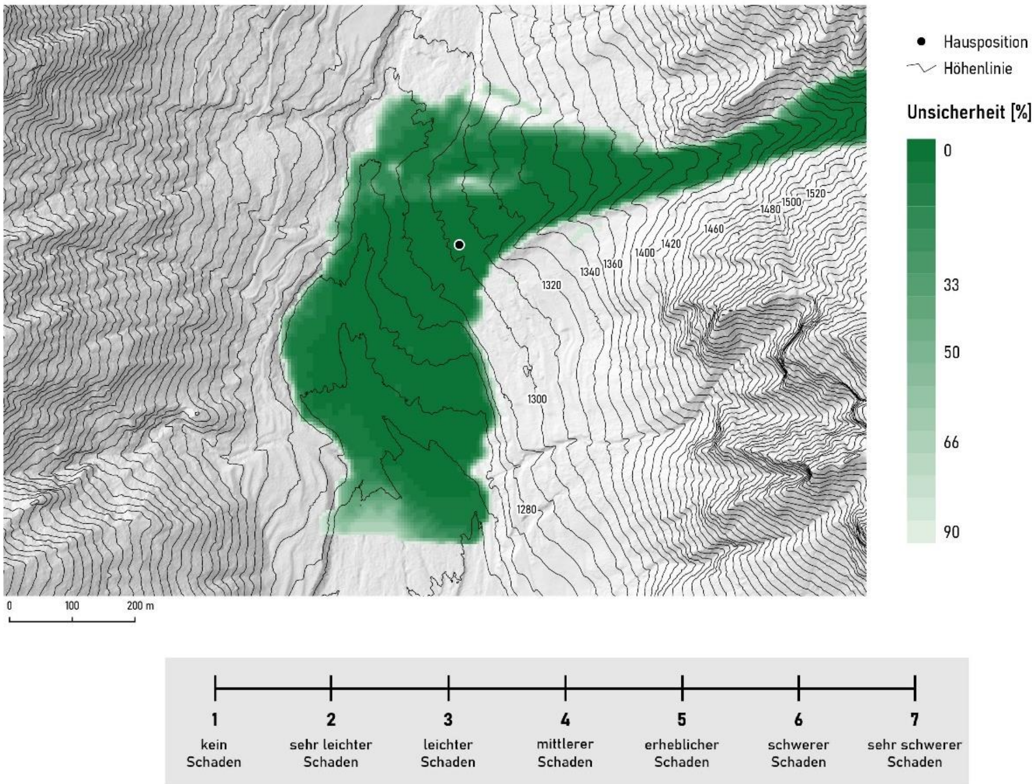
Deine Schadenseinschätzung wirst Du mit einem Mausklick auf die entsprechende Ziffer (1-7) in der Schadensskala unterhalb der Karte abgeben. Du kannst nur ganzzahlige Schadenseinschätzungen abgeben. Dieser Mausklick bringt Dich jeweils zur nächsten Einschätzungsaufgabe. Insgesamt wirst du 40 Einschätzungen abgeben. Bevor die Aufgaben starten, wirst Du 2 Übungsaufgaben lösen, um Dich mit der Aufgabenstellung vertraut zu machen.

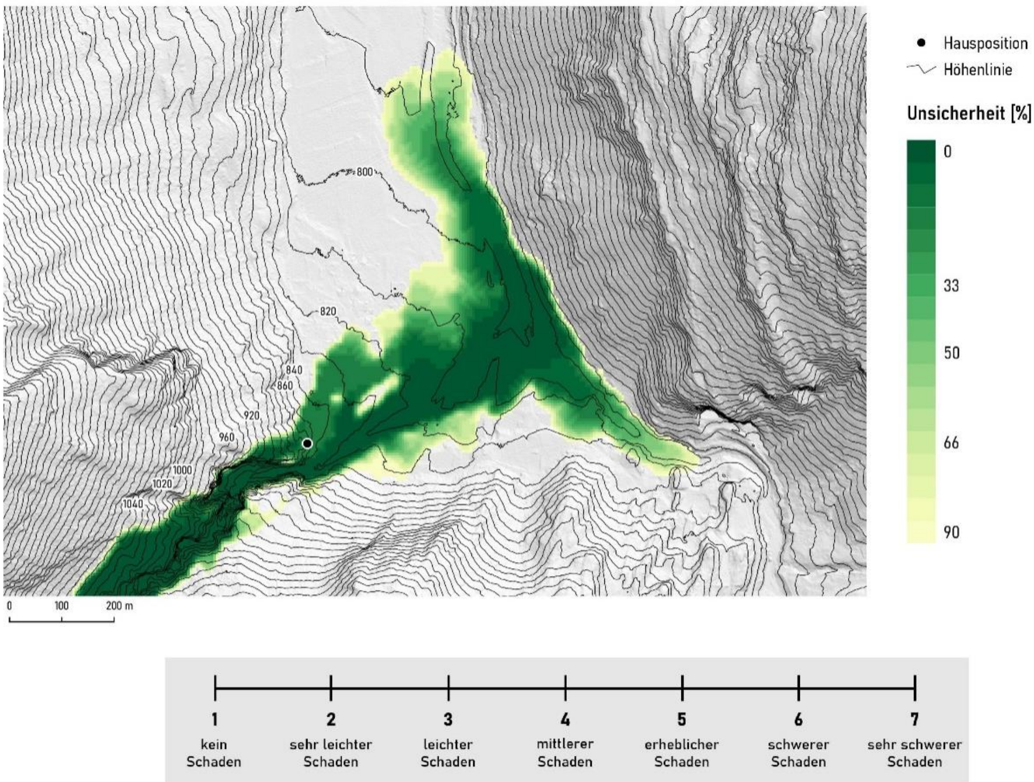


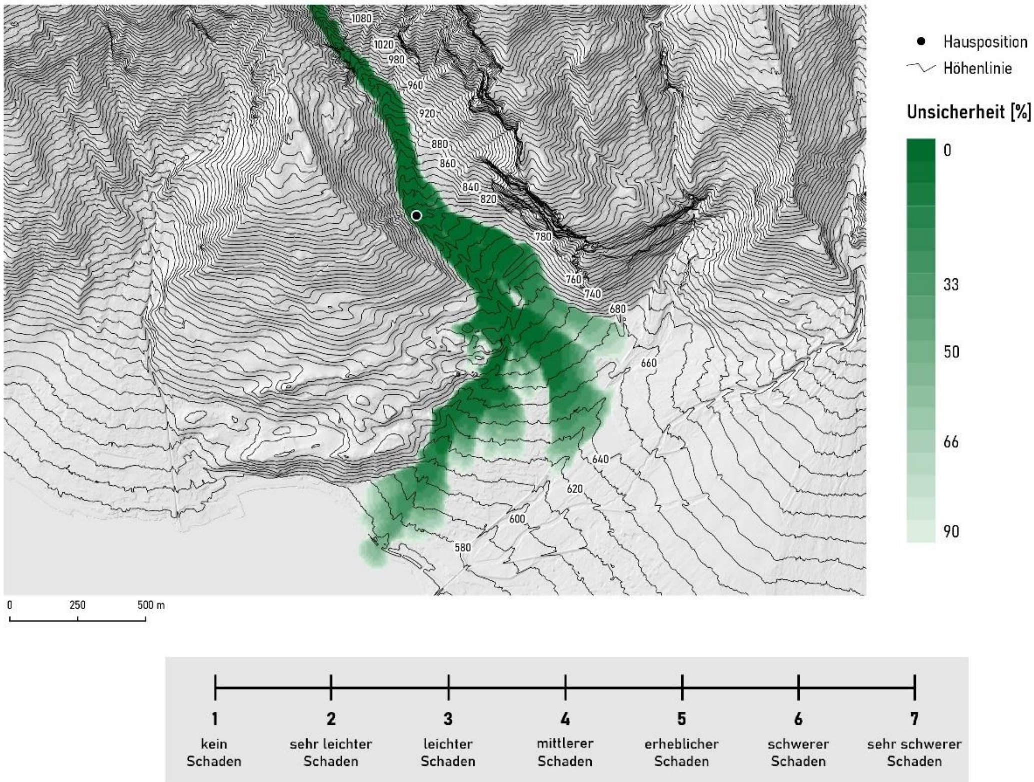
Falls du hierzu noch Fragen hast, darfst du dich gerne bei der Experimentleiterin melden.

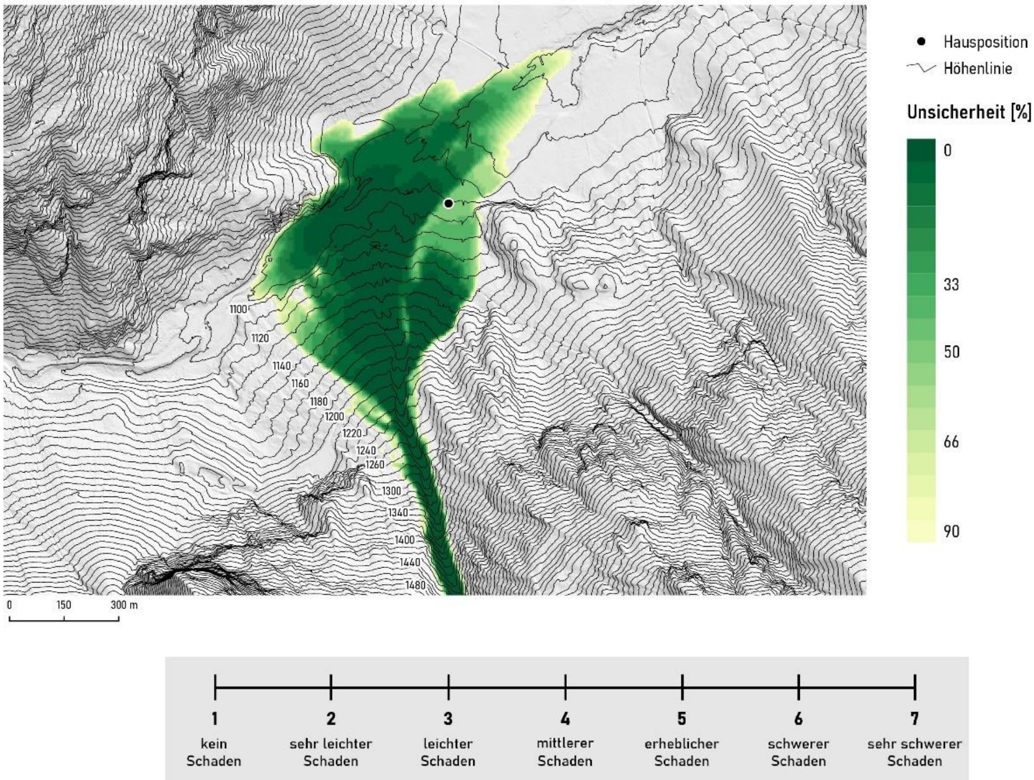
P. Map Stimuli

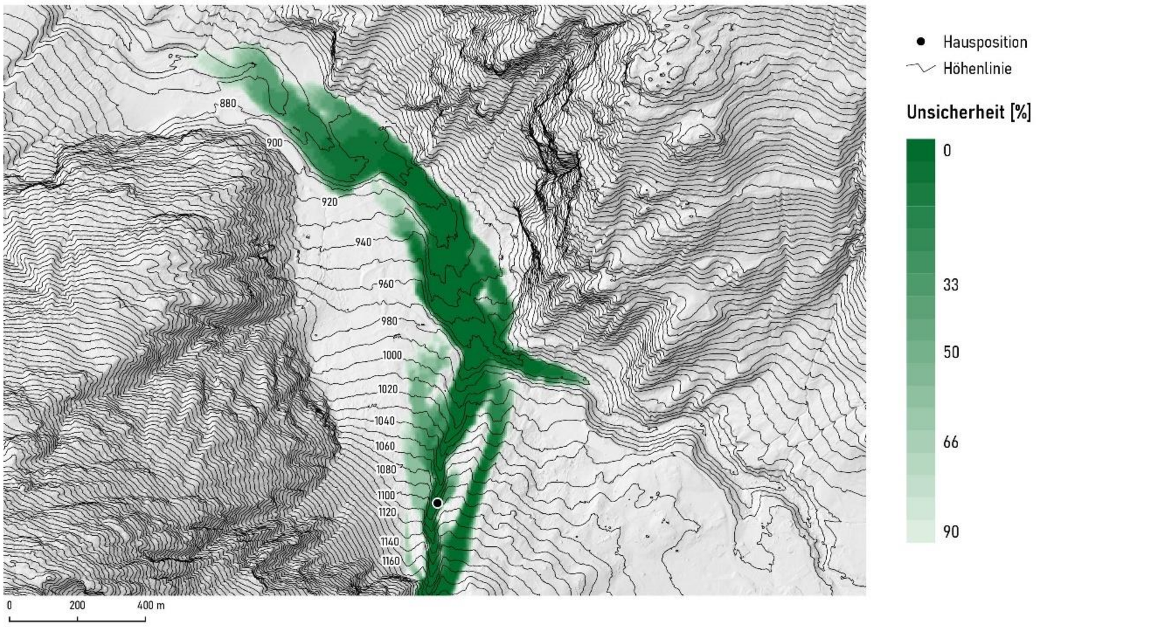
A selection of 10 maps (one per debris flow location) is displayed below to give an impression of the tested map stimuli.



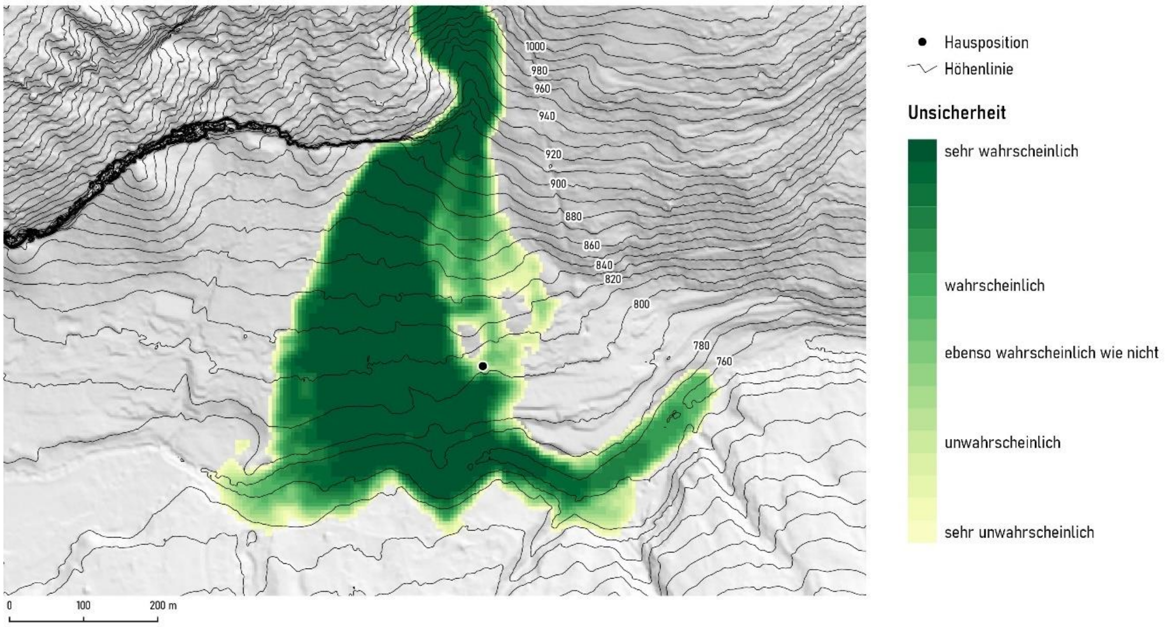


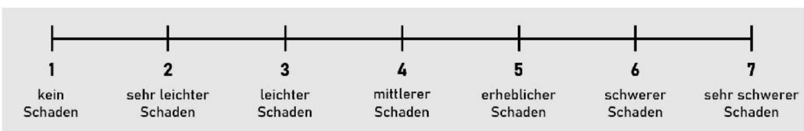

Hausposition
Höhenlinie
Unsicherheit [%]
0
33
50
66
90
0
250
500 m
1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden

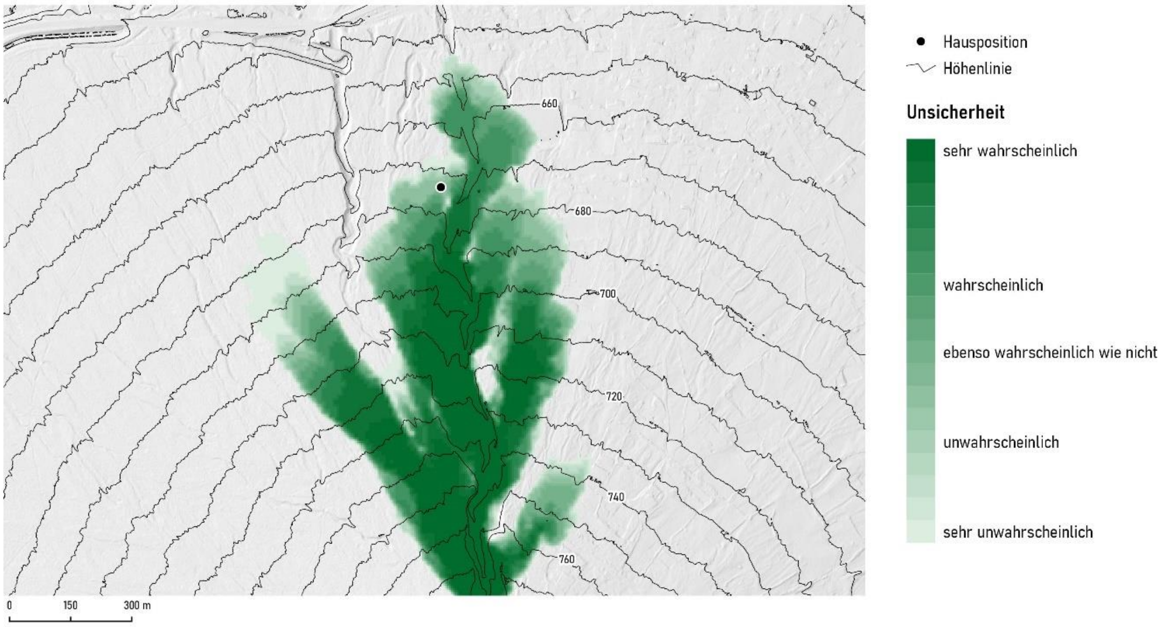

Hausposition
Höhenlinie
Unsicherheit [%]
0
33
50
66
90
0
150
300 m
1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden


880
900
920
940
960
980
1000
1020
1040
1060
1080
1100
1120
1140
1160
0 200 400 m
Hausposition
Höhenlinie
Unsicherheit [%]
0
33
50
66
90

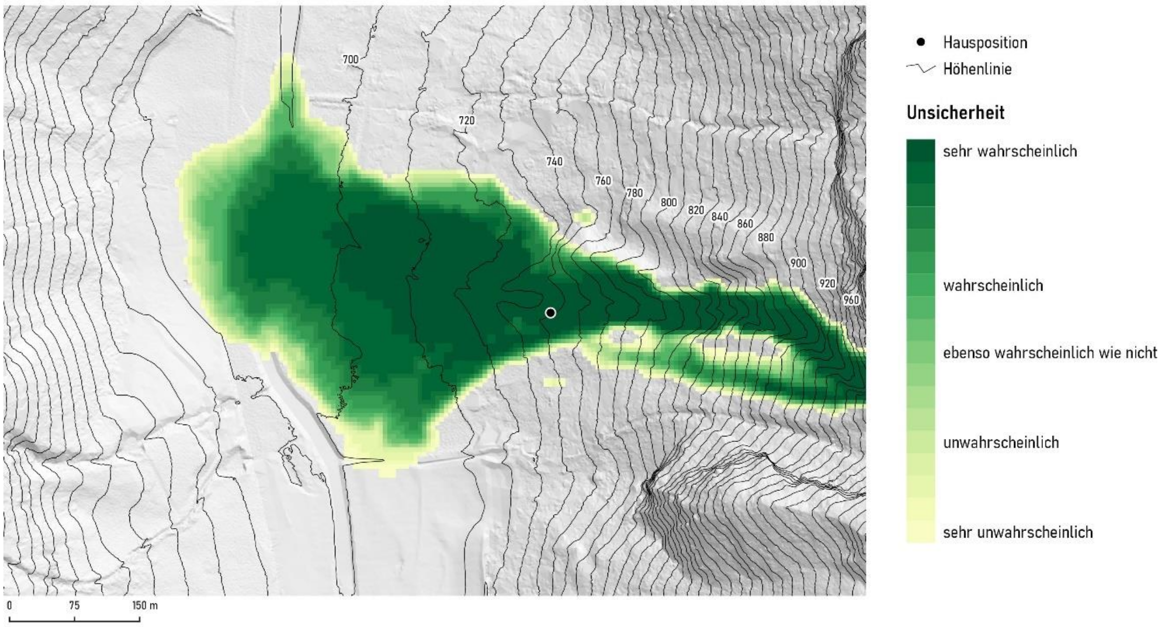

1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden


1000
980
960
940
920
900
880
860
840
820
800
780
760
0 100 200 m
Hausposition
Höhenlinie
Unsicherheit
sehr wahrscheinlich
wahrscheinlich
ebenso wahrscheinlich wie nicht
unwahrscheinlich
sehr unwahrscheinlich

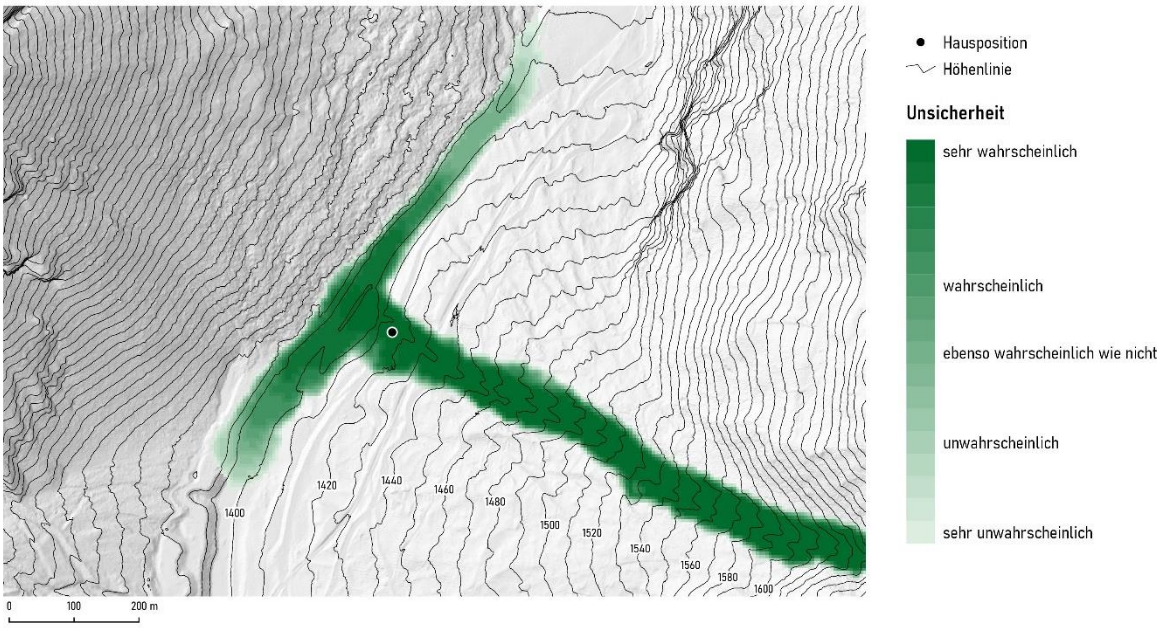

1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden


660
680
700
720
740
760
0
150
300 m
Hausposition
Höhenlinie
Unsicherheit
sehr wahrscheinlich
wahrscheinlich
ebenso wahrscheinlich wie nicht
unwahrscheinlich
sehr unwahrscheinlich

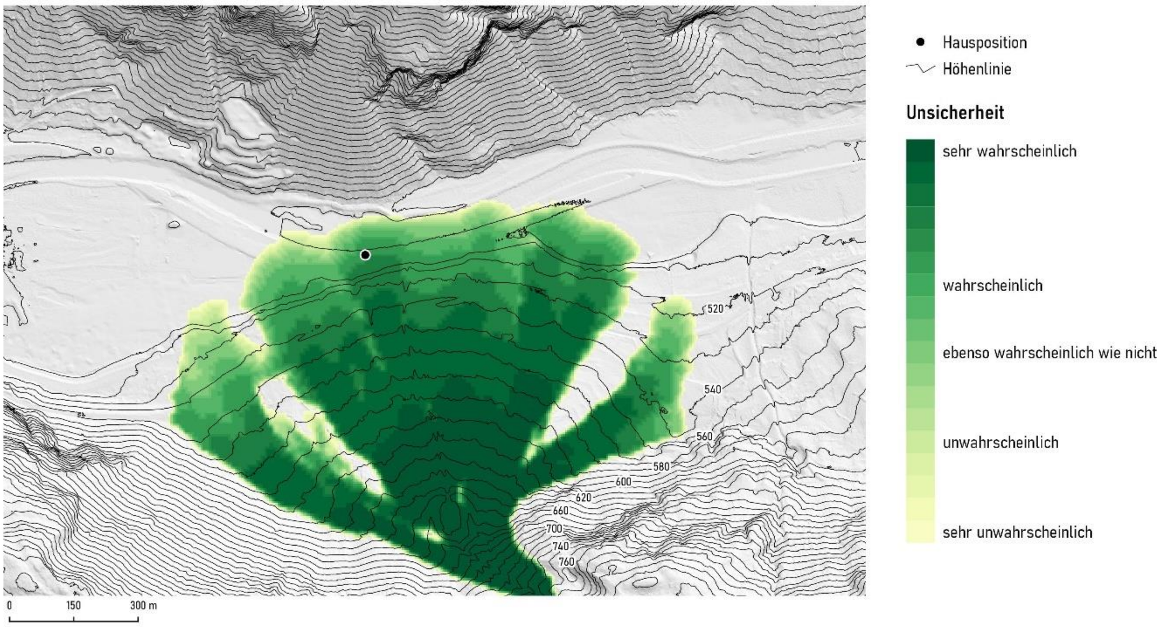

1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden

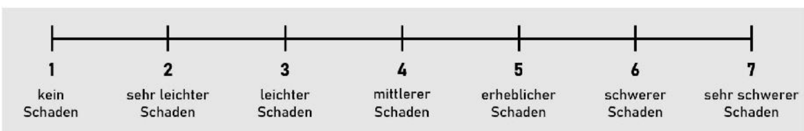

700
720
740
760
780
800
820
840
860
880
900
920
960
0
75
150 m
Hausposition
Höhenlinie
Unsicherheit
sehr wahrscheinlich
wahrscheinlich
ebenso wahrscheinlich wie nicht
unwahrscheinlich
sehr unwahrscheinlich


1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden


Hausposition
Höhenlinie
Unsicherheit
sehr wahrscheinlich
wahrscheinlich
ebenso wahrscheinlich wie nicht
unwahrscheinlich
sehr unwahrscheinlich
1400
1420
1440
1460
1480
1500
1520
1540
1560
1580
1600
0
100
200 m


1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden


Hausposition
Höhenlinie
Unsicherheit
sehr wahrscheinlich
wahrscheinlich
ebenso wahrscheinlich wie nicht
unwahrscheinlich
sehr unwahrscheinlich
520
540
560
580
600
620
660
700
740
760
0
150
300 m


1
kein Schaden
2
sehr leichter Schaden
3
leichter Schaden
4
mittlerer Schaden
5
erheblicher Schaden
6
schwerer Schaden
7
sehr schwerer Schaden

Q. Post-Test Questionnaire

Master Arbeit: Visualisierung und Kommunikation von Unsicherheiten bei Murgängen - Fragebogen 2

Zwischengespeicherte Umfrage laden

0%

Master Arbeit: Visualisierung und Kommunikation von Unsicherheiten bei Murgängen - Fragebogen 2

Den grössten Teil der Studie hast Du hinter Dir. Nun werden Dir noch einige Fragen zum Aufgabenblock gestellt.

Weiter

Teil 4

* Diese Frage wird von der Experimentleiterin ausgefüllt!

Teilnehmer/innen-ID:

- Deine Antwort muss mindestens 0 sein.
- In diesem Feld kann nur ein ganzzahliger Wert eingetragen werden.

Teil 5

* Beschreibe kurz dein Vorgehen beim Lösen deiner Entscheidungsaufgabe. Hast du eine bestimmte Strategie verfolgt? Wenn ja, welche?

*** Welche Elemente der Kartendarstellung hast du benutzt, um deine Entscheidung zu treffen?**

Bitte wähle die zutreffenden Punkte aus.

- ☐ Farbschema
- ☐ Kartenlegende
- ☐ Massstab
- ☐ Hausposition
- ☐ Höhenlinie
- ☐ Schummerung (Kartenhintergrund)

Gab es noch weitere Faktoren, welche du in deiner Entscheidung berücksichtigt hast? Wenn ja, bitte notiere sie hier:

*** Wie wichtig war die Information der Höhenlinen bei deiner Entscheidung? Begründe deine Antwort im Kommentarfeld rechts.**

Bitte wähle eine der folgenden Antworten:

- ○ 1 - gar nicht wichtig
- ○ 2
- ○ 3 - neutral
- ○ 4
- ○ 5 - sehr wichtig

Bitte gib hier Deinen Kommentar ein:

*** Wie wichtig war die Information in der Kartenlegende (rechts von der Karte) zur Entscheidungsfindung? Begründe deine Antwort im Kommentarfeld rechts.**

Bitte wähle eine der folgenden Antworten:

- ○ 1 - gar nicht wichtig
- ○ 2
- ○ 3 - neutral
- ○ 4
- ○ 5 - sehr wichtig

Bitte gib hier Deinen Kommentar ein:

*** Wie viel Vertrauen hast du in die Murgangvorhersage gesetzt?**

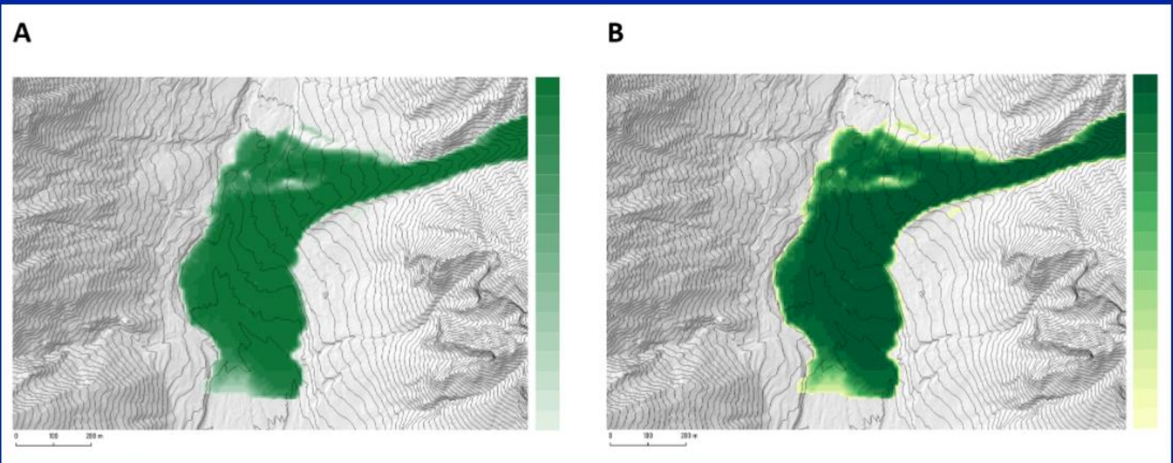
Bitte wähle eine der folgenden Antworten:

- ○ 1 - sehr wenig Vertrauen
- ○ 2
- ○ 3 - mittleres Vertrauen
- ○ 4
- ○ 5 - sehr viel Vertrauen

Möchtest du noch etwas bezüglich deines Entscheidungsprozesses im Aufgabenblock ergänzen?

Teil 6

*** Auf einer Skala von 1 bis 5*, wie hilfreich waren die beiden Visualisierungsmethoden beim Treffen der Entscheidung?**

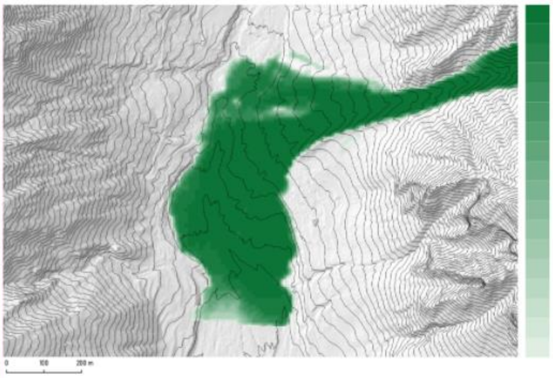


	1	2	3	4	5
Farbschema A (hellgrün - dunkelgrün)	○	○	○	○	○
Farbschema B (hellgelb - dunkelgrün)	○	○	○	○	○

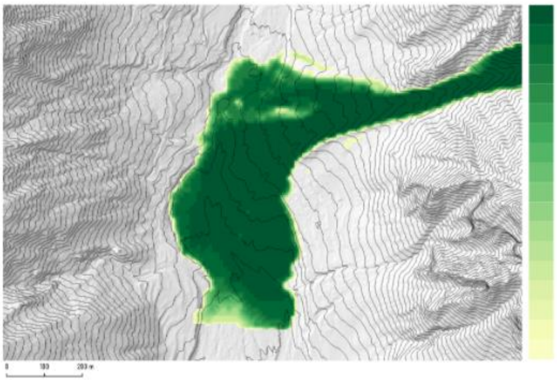
* Ein Wert von 1 bedeutet, die Visualisierung war gar nicht hilfreich. Ein Wert von 5 bedeutet, dass dir die Visualisierung sehr hilfreich war.

* Erläutere, was an den beiden Visualisierungsmethoden hilfreich war, und was dir die Entscheidung erschwert hat?

A

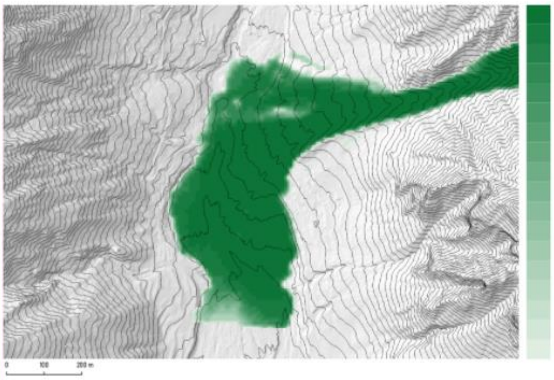


B

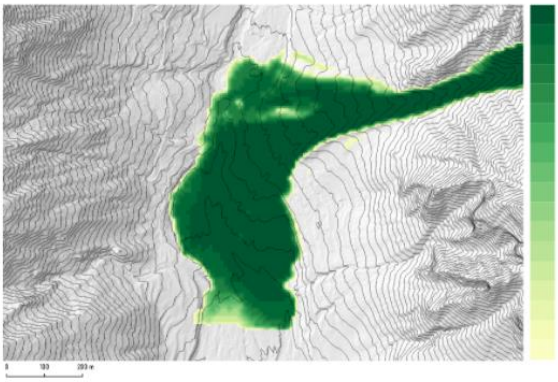


* Auf einer Skala von 1 bis 5*, wie gut gefallen dir die Visualisierungsmethoden optisch?

A



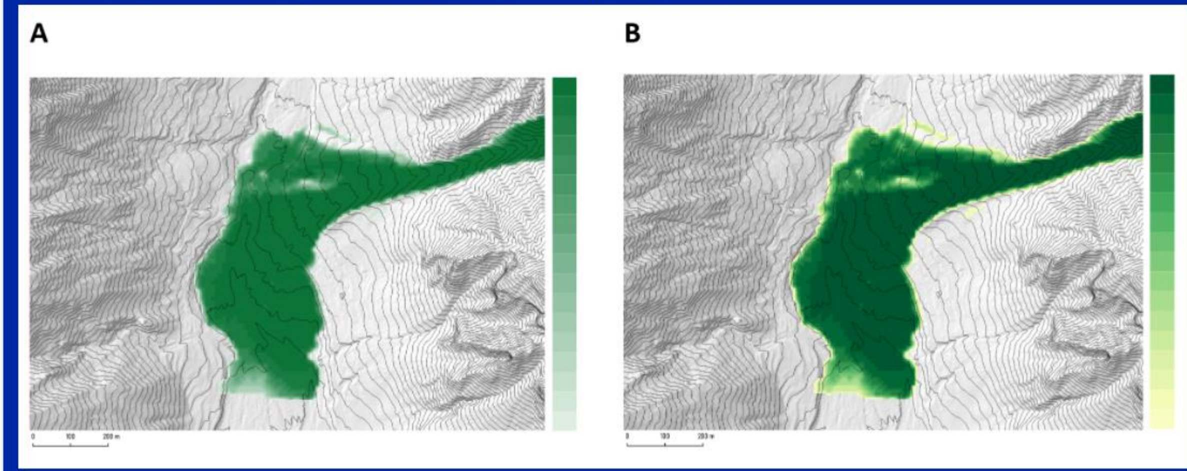
B



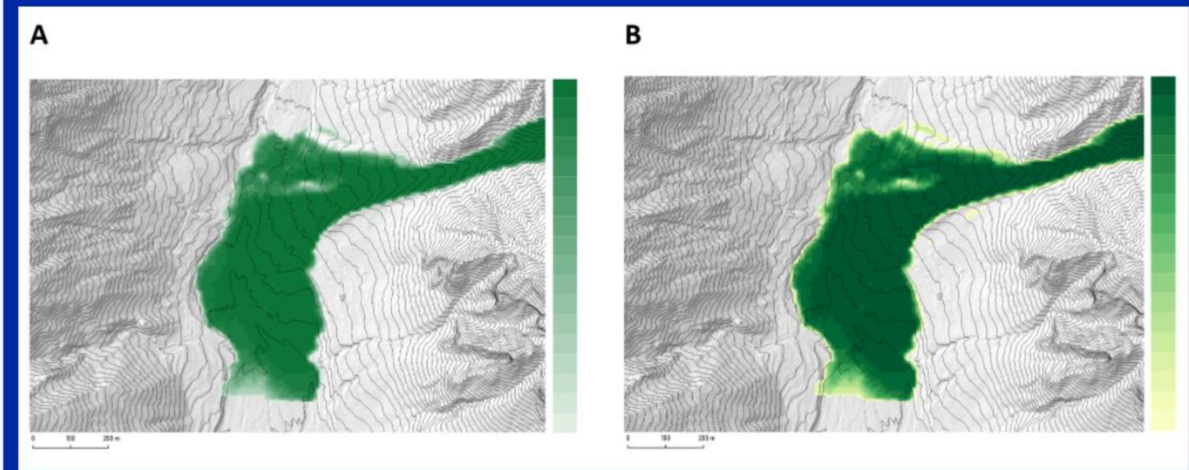
	1	2	3	4	5
Farbschema A (hellgrün - dunkelgrün)	○	○	○	○	○
Farbschema B (hellgelb - dunkelgrün)	○	○	○	○	○

* Ein Wert von 1 bedeutet, die Visualisierung hat dir gar nicht gefallen. Ein Wert von 5 bedeutet, dass dir die Visualisierung sehr gut gefallen hat.

* Erläutere, was dir an den beiden Visualisierungsmethoden optisch gefallen hat und was nicht?



Möchtest du noch etwas zu den beiden Visualisierungsmethoden ergänzen?



Teil 7

*** Wie verständlich war das Szenario im Aufgabenblock für dich?**

Bitte wähle eine der folgenden Antworten:

- ○ 1 - unverständlich
- ○ 2
- ○ 3 - neutral
- ○ 4
- ○ 5 - verständlich

*** Wie schwierig war das Lösen der Aufgaben für dich?**

Bitte wähle eine der folgenden Antworten:

- ○ 1 - nicht schwierig
- ○ 2
- ○ 3 - neutral
- ○ 4
- ○ 5 - sehr schwierig

Falls die Aufgabe Schwierigkeiten bot, welche waren das?

*** Denkst du, die Darstellung der Unsicherheit hat deine Entscheidung unterstützt oder erschwert?**

Bitte wähle eine der folgenden Antworten:

- ○ 1 - erschwert
- ○ 2
- ○ 3 - kein Einfluss
- ○ 4
- ○ 5 - unterstützt

* Erläutere, inwiefern die Darstellung der Unsicherheit deine Entscheidung unterstützt oder erschwert hat.

Teil 8

* Hast du einen der dargestellten Murgang-Standorte / -Graben gekannt? Wenn ja, schreibe den Orts- / Grabennamen ins Kommentarfeld rechts.

Bitte wähle eine der folgenden Antworten:

- Ja
- Nein

Bitte gib hier Deinen Kommentar ein:

Teil 9

* Wie würdest du die Darstellungen im Experiment verändern, um die Entscheidungen weiter zu erleichtern?

Hast du noch Anmerkungen oder Kommentare zur Studie?

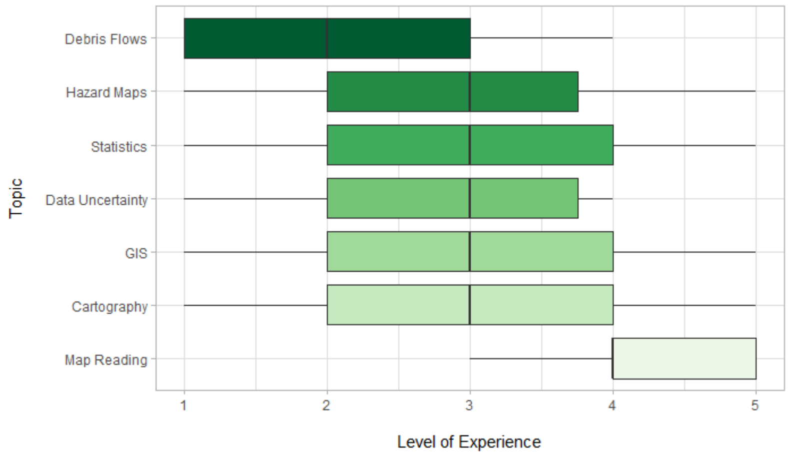
R. Additional Results

Pre-Test Questionnaire

Table 32: Results of self-assessment of experience in different topics for the numerical and the verbal group.

	Numerical		Verbal	
Topic	**Mean**	**SD**	**Mean**	**SD**
Map reading	4.09	0.75	3.81	0.98
Cartography	3.09	1.19	2.76	1.61
GIS	2.95	1.40	2.86	1.85
Data uncertainty	2.64	1.09	2.90	1.41
Statistics	2.95	1.17	2.95	1.12
Hazard maps	2.86	1.08	2.43	1.16
Debris flows	2.27	1.08	1.86	1.01

A – numerical



B – verbal

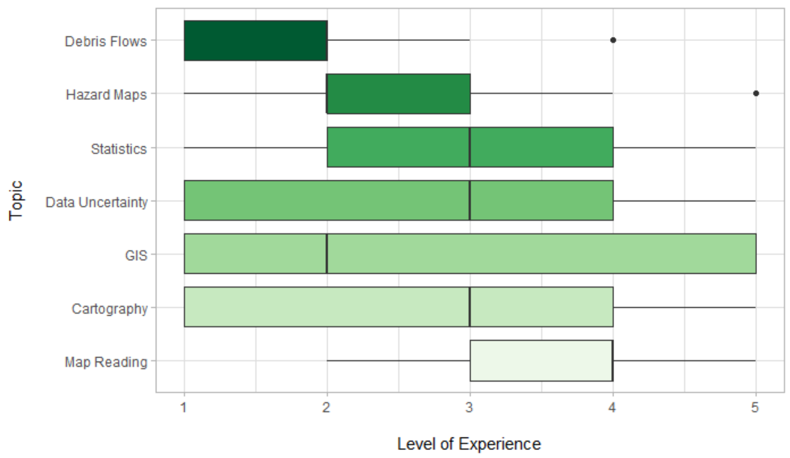
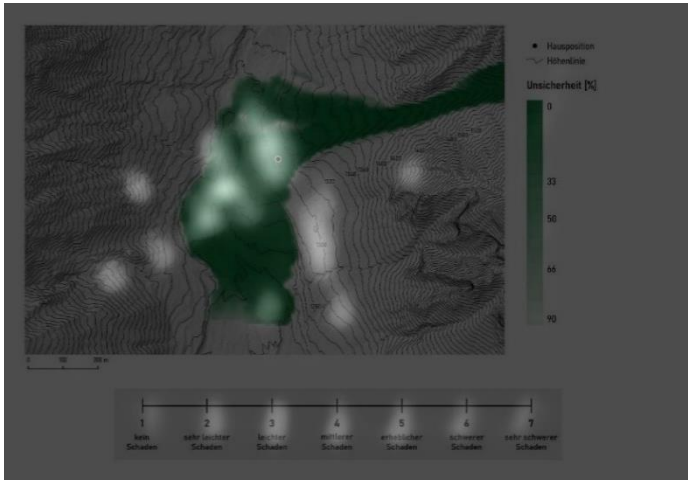


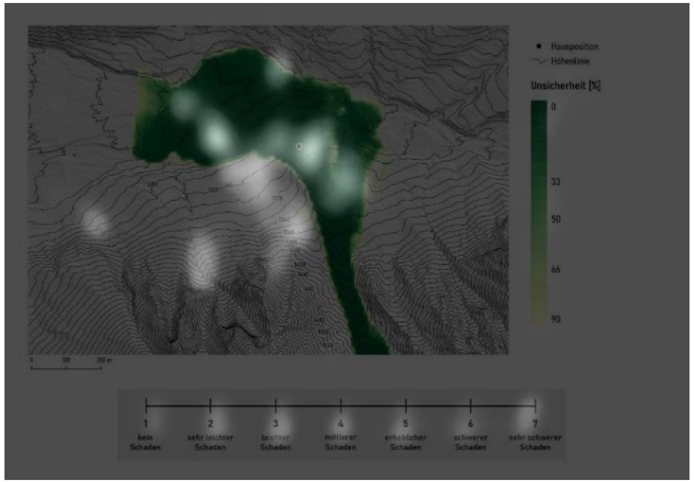
*Figure 65: Boxplot of self-assessment of experience regarding different topics for the numerical (A) and verbal group (B). The horizontal line represents the median, the box shows the 25th and the 75th percentile, while the whiskers indicate 1.5 * the interquartile range.*

RQ1: Uncertainty Visualization

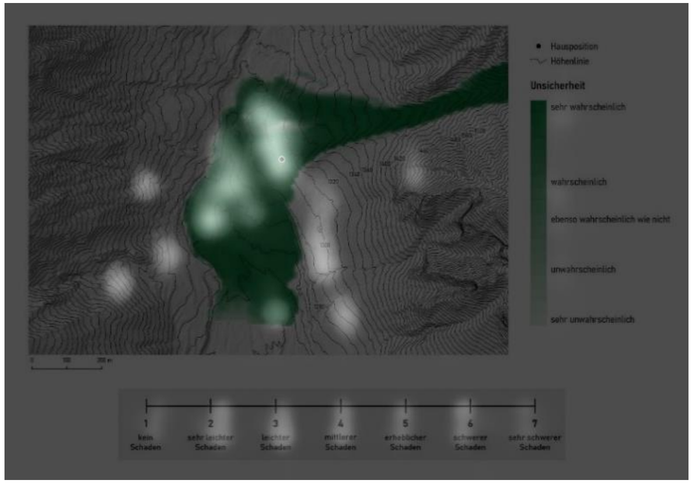
A – numerical / single-hue



B – numerical /multi-hue



C – verbal / single-hue



D – verbal / multi-hue

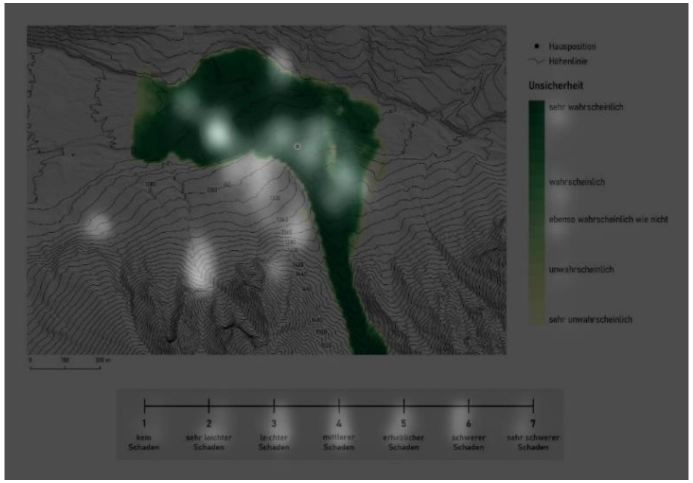


Figure 66: Heatmaps of relative duration of fixations per communication group for the two visualization methods. Bright areas represent regions which were attended longer than darker, more obscured areas.

Table 33: Overview of positive and negative aspects of the colour schemes in terms of their suitability for decision-making, total count of mentions and their distribution between the communication groups (N = numerical, V = verbal).

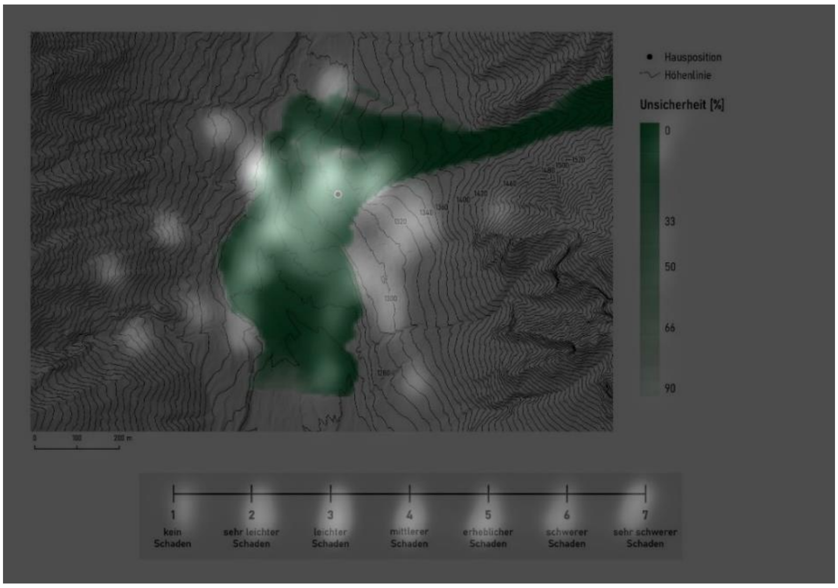
Colour Scheme	Pro	Total	N	V	Contra	Total	N	V
Single-hue	Uncertainty evoking	3	2	1	Brightness levels hard to distinguish	19	7	11
	Intuitiveness	2	2	-	Low contrast from background	8	2	6
	Linearity of colour scheme	1	1	-	Bright values hard to distinguish	6	4	2
	Quick overview	1	1	-	Hill shade disguised	1	1	-
	Contours stuck out	1	-	1				
Multi-hue	Clear colour levels	25	12	13	Saliency of high uncertainty values	4	2	2
	Judgment of high uncertainty regions	10	5	5	Hill shade disguised	1	1	-
	Easy matching of map & uncertainty	7	6	1	Quick transition from green to yellow	1	-	1
	High contrast from background	4	1	3	Too many colour levels	1	1	-
	Intuitiveness	1	-	1	Contours hard to see	1	-	1
					Yellow triggers fear	1	-	1

Table 34: Overview of positive and negative aspects of the colour schemes in terms of their aesthetics, total count of mentions and their distribution between the communication groups (N = numerical, V = verbal).

Colour Scheme	**Pro**	**Total**	**N**	**V**	**Contra**	**Total**	**N**	**V**
Single-hue	Calm	5	1	4	Green ≠ danger / risk	5	3	2
	Intuitive	3	2	1	Bland	3	1	2
	High contrast from background	2	2		Hard to interpret transition in value	1	1	
	Aesthetic	2	1	1	Less aesthetic	1	1	
	Simple	1	1		Less interesting	1		1
	Harmony with background	1		1				
	No sharp border	1		1				
Multi-hue	Aesthetic	6	4	2	Green ≠ danger / risk	5	3	2
	Easy to interpret	4	2	2	High saliency yellow	1	1	
	High saliency	4	1	3	Less aesthetic	1	1	
	Multiple hues	2	1	1	Border too sharp	1		1
	High contrast in hue	2		2				
	High contrast from background	2	2					
	Transition of hues	2		2				
	Seems 3D	2		2				
	Colour blind friendly	1		1				
	Warm hues	1	1					
	Intuitive	1	1					
	More detailed	1		1				

RQ2: Uncertainty Communication

A – numerical



B – verbal

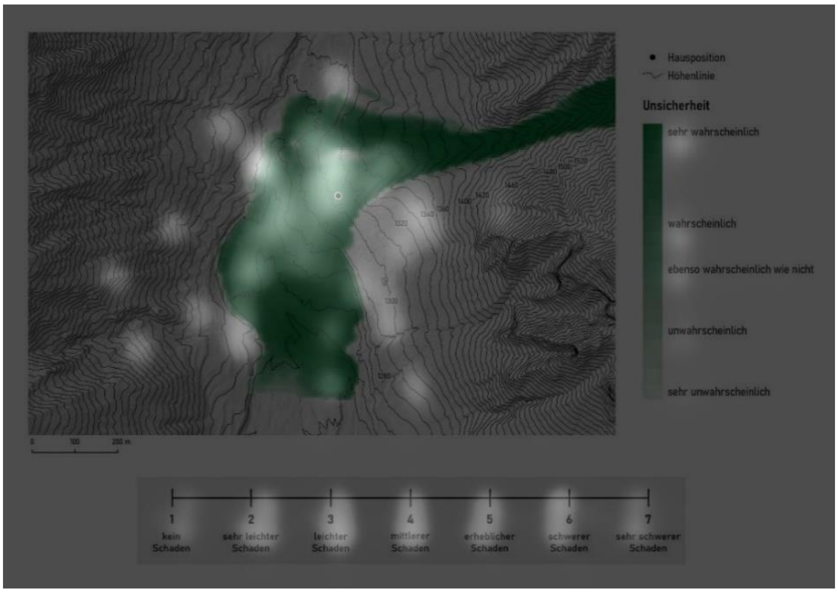


Figure 67: Heatmaps of relative duration of fixations per communication group. Bright areas represent regions which were attended longer than darker, more obscured areas.

Table 35: Overview of the positive and negative aspects regarding the map legend, total count of mentions and their distribution between the communication groups (N = numerical, V = verbal).

Pro	Total	N	V	Contra	Total	N	V
Matching colour to legend label	17	9	8	Difficulty matching colour to legend label	2	2	
Damage estimation	6	4	2	Other factors more important	2		2
Define affected area	5	3	2	Only taken into account during first few trials	2	1	1
Gain overview	5	1	4	Uncertainty definition unclear	1	1	
Debris flow volume estimation	1		1	Too many colour levels	1	1	

RQ1 & RQ2: Interaction between Visualization & Communication Method

*Table 36: Contrasts of the ARTool output for the interaction between communication and visualization method regarding the response time (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

contrast	estimate	SE	df	t.ratio	p.value	sig
numerical,multi_hue - numerical,single_hue	-34.80	28.1	1675	-1.24	0.60	
numerical,multi_hue - verbal,multi_hue	-141.70	87.4	45.7	-1.62	0.38	
numerical,multi_hue - verbal,single_hue	-79.30	87.4	45.7	-0.91	0.80	
numerical,single_hue - verbal,multi_hue	-106.80	87.4	45.7	-1.22	0.62	
numerical,single_hue - verbal,single_hue	-44.40	87.4	45.7	-0.51	0.96	
verbal,multi_hue - verbal,single_hue	62.40	28.8	1675	2.17	0.13	

RQ3: Additional Information

Table 37: Descriptive statistics on the parameters slope, distance and uncertainty value at the house locations.

Parameter	Normality	Mean	Median	SD	Min	Max
Slope [°]	p-value = 0.400	21.89	19.77	13.90	3.61	52.85
Distance [m]	p-value = 0.004	95.29	54.35	97.16	5.62	344.71
Uncertainty Value [%]	p-value = 0.014	40.81	35.01	37.56	0	100

*Table 38: ARTool output for the effect of the uncertainty value category on damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

	F	Df	Df.res	Pr(>F)	sig
Uncertainty value	2210.80	2	1675	< 2.22e-16	***

Table 39: Contrasts of the ARTool output for the effect of the uncertainty value category on damage estimate.

contrast	estimate	SE	df	t.ratio	p.value	d
high - low	-940.00	14.50	1675	-64.82	<.0001	-3.68
high - medium	-353.00	16.70	1675	-21.19	<.0001	-1.38
low - medium	587.00	15.30	1675	38.24	<.0001	2.30

*Table 40: ARTool output for the effect of the slope category on damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

	F	Df	Df.res	Pr(>F)	sig
Slope	14.22	2	1675	7.55E-07	***

Table 41: Contrasts of the ARTool output for the effect of the slope category on damage estimate.

contrast	estimate	SE	df	t.ratio	p.value	d
high - low	153.90	30.00	1675	5.12	<.0001	0.32
high - medium	41.90	28.10	1675	1.49	0.30	0.09
low - medium	-112.00	28.10	1675	-3.99	0.0002	-0.23

*Table 42: ARTool output for the effect of the distance category on damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

	F	Df	Df.res	Pr(>F)	sig
Distance	1643.80	2	1675	< 2.22e-16	***

Table 43: Contrasts of the ARTool output for the effect of the distance category on damage estimate.

contrast	estimate	SE	df	t.ratio	p.value	d
high - low	-1004.00	17.60	1675	-57.03	<.0001	-3.55
high - medium	-419.00	16.50	1675	-25.45	<.0001	-1.48
low - medium	585.00	16.50	1675	35.51	<.0001	2.07

*Table 44: ARTool output for the interaction between slope and distance category in terms of the damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

	F	Df	Df.res	Pr(>F)	sig
Slope	77.31	2	1669	< 2.22e-16	***
Distance	2237.45	2	1669	< 2.22e-16	***
Slope : Distance	154.65	4	1669	< 2.22e-16	***

*Table 45: Contrasts of the ARTool output for the interaction between slope and distance category in terms of the damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

contrast	estimate	SE	df	t.ratio	p.value	sig.
high,high - high,low	-1152.76	25.40	1669	-45.38	<.0001	***
high,high - high,medium	-882.49	25.40	1669	-34.74	<.0001	***
high,high - low,high	-33.78	25.40	1669	-1.33	0.92	
high,high - low,low	-1124.35	25.40	1669	-44.26	<.0001	***
high,high - low,medium	-415.37	25.40	1669	-16.35	<.0001	***
high,high - medium,high	-406.54	25.40	1669	-16.01	<.0001	***
high,high - medium,low	-1174.21	25.40	1669	-46.23	<.0001	***
high,high - medium,medium	-482.62	22.00	1669	-21.94	<.0001	***
high,low - high,medium	270.28	25.40	1669	10.64	<.0001	***
high,low - low,high	1118.98	25.40	1669	44.05	<.0001	***
high,low - low,low	28.42	25.40	1669	1.12	0.97	
high,low - low,medium	737.40	25.40	1669	29.03	<.0001	***
high,low - medium,high	746.22	25.40	1669	29.38	<.0001	***
high,low - medium,low	-21.44	25.40	1669	-0.84	1.00	
high,low - medium,medium	670.14	22.00	1669	30.46	<.0001	***
high,medium - low,high	848.70	25.40	1669	33.41	<.0001	***
high,medium - low,low	-241.86	25.40	1669	-9.52	<.0001	***
high,medium - low,medium	467.12	25.40	1669	18.39	<.0001	***
high,medium - medium,high	475.94	25.40	1669	18.74	<.0001	***
high,medium - medium,low	-291.72	25.40	1669	-11.49	<.0001	***
high,medium - medium,medium	399.86	22.00	1669	18.18	<.0001	***
low,high - low,low	-1090.57	25.40	1669	-42.93	<.0001	***
low,high - low,medium	-381.58	25.40	1669	-15.02	<.0001	***
low,high - medium,high	-372.76	25.40	1669	-14.68	<.0001	***
low,high - medium,low	-1140.42	25.40	1669	-44.90	<.0001	***
low,high - medium,medium	-448.84	22.00	1669	-20.40	<.0001	***
low,low - low,medium	708.98	25.40	1669	27.91	<.0001	***
low,low - medium,high	717.81	25.40	1669	28.26	<.0001	***
low,low - medium,low	-49.86	25.40	1669	-1.96	0.57	
low,low - medium,medium	641.73	22.00	1669	29.17	<.0001	***
low,medium - medium,high	8.82	25.40	1669	0.35	1.00	
low,medium - medium,low	-758.84	25.40	1669	-29.87	<.0001	***
low,medium - medium,medium	-67.25	22.00	1669	-3.06	0.06	
medium,high - medium,low	-767.66	25.40	1669	-30.22	<.0001	***
medium,high - medium,medium	-76.08	22.00	1669	-3.46	0.0162 *	
medium,low - medium,medium	691.59	22.00	1669	31.44	<.0001	***

*Table 46: ARTool output for the interaction between slope and uncertainty value category in terms of the damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

	F	Df	Df.res	Pr(>F)	sig
Slope	35.6657	2	1670	6.80E-16	***
Uncertainty value	2029.448	2	1670	< 2.22e-16	***
Slope : Uncertainty value	9.1736	3	1670	5.08E-06	***

*Table 47: Contrasts of the ARTool output for the interaction between slope and uncertainty value category in terms of the damage estimate (sig: 0, ***, 0.001, **, 0.01, *, 0.05).*

contrast	estimate	SE	df	t.ratio	p.value	sig
high,high - high,low	-1017.62	23.20	1670	-43.87	<.0001	***
high,high - low,high	-180.59	24.40	1670	-7.39	<.0001	***
high,high - low,low	-1124.35	26.80	1670	-41.98	<.0001	***
high,high - low,medium	-356.53	32.80	1670	-10.87	<.0001	***
high,high - medium,high	-84.86	32.80	1670	-2.59	0.16	
high,high - medium,low	-1025.55	24.40	1670	-41.95	<.0001	***
high,high - medium,medium	-482.62	23.20	1670	-20.81	<.0001	***
high,low - low,high	837.03	20.50	1670	40.92	<.0001	***
high,low - low,low	-106.72	23.20	1670	-4.60	0.0001	***
high,low - low,medium	661.10	29.90	1670	22.08	<.0001	***
high,low - medium,high	932.76	29.90	1670	31.15	<.0001	***
high,low - medium,low	-7.92	20.50	1670	-0.39	1.00	
high,low - medium,medium	535.00	18.90	1670	28.25	<.0001	***
low,high - low,low	-943.76	24.40	1670	-38.60	<.0001	***
low,high - low,medium	-175.94	30.90	1670	-5.69	<.0001	***
low,high - medium,high	95.73	30.90	1670	3.10	0.04	*
low,high - medium,low	-844.96	21.90	1670	-38.64	<.0001	***
low,high - medium,medium	-302.03	20.50	1670	-14.77	<.0001	***
low,low - low,medium	767.82	32.80	1670	23.41	<.0001	***
low,low - medium,high	1039.49	32.80	1670	31.69	<.0001	***
low,low - medium,low	98.80	24.40	1670	4.04	0.0014	**
low,low - medium,medium	641.73	23.20	1670	27.67	<.0001	***
low,medium - medium,high	271.66	37.90	1670	7.17	<.0001	***
low,medium - medium,low	-669.02	30.90	1670	-21.63	<.0001	***
low,medium - medium,medium	-126.09	29.90	1670	-4.21	0.0007	***
medium,high - medium,low	-940.68	30.90	1670	-30.42	<.0001	***
medium,high - medium,medium	-397.76	29.90	1670	-13.28	<.0001	***
medium,low - medium,medium	542.92	20.50	1670	26.54	<.0001	***

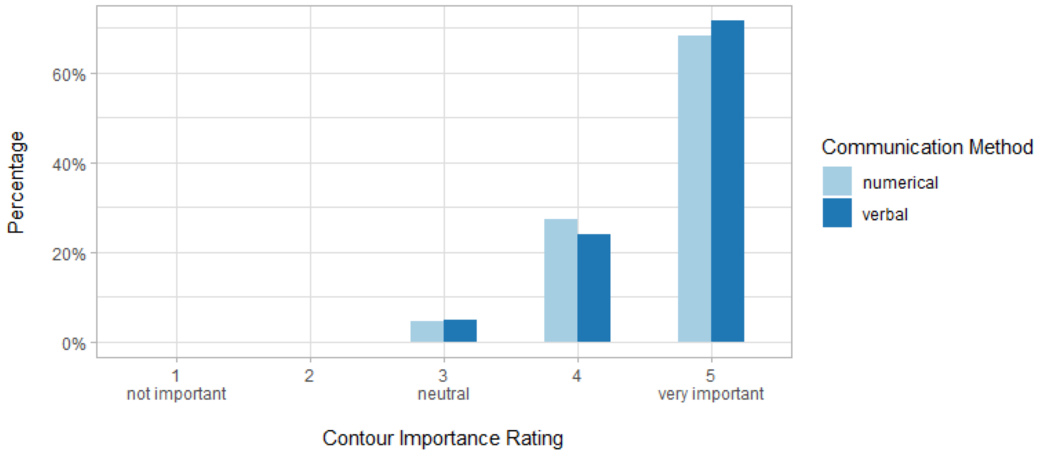


Figure 68: Relative distribution of contour line importance rating for the two communication methods.

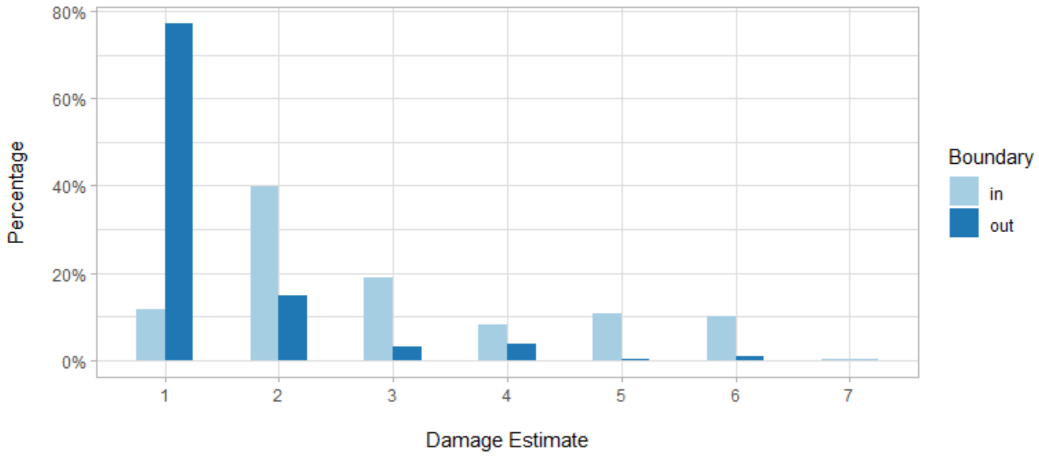


Figure 69: Relative distribution of damage estimates at high-distance locations inside and outside the debris flow boundary.

Post-Test Questionnaire

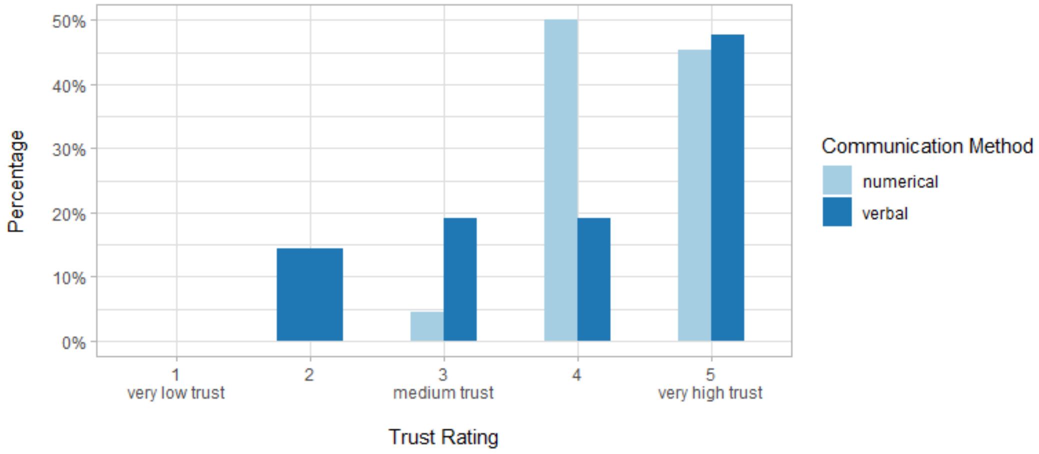


Figure 70: Relative distribution of trust rating regarding the uncertainty information for the two communication methods.

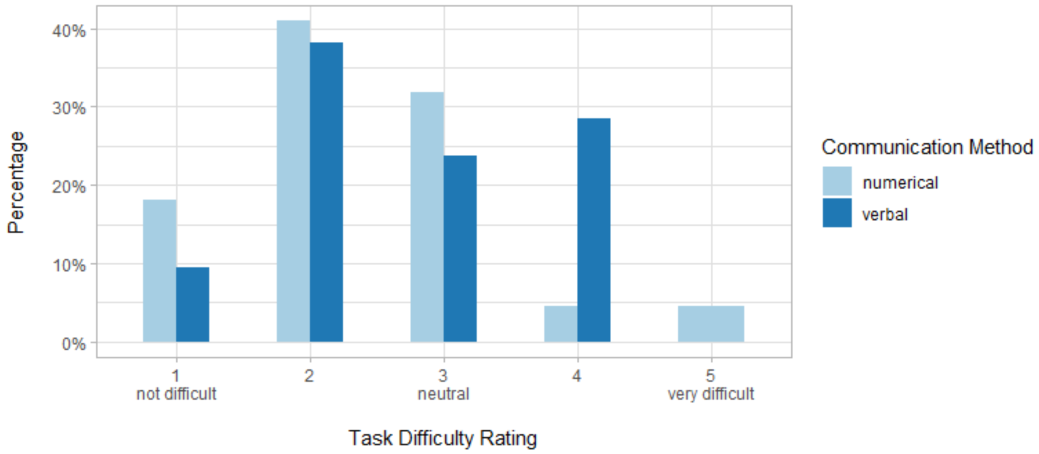


Figure 71: Relative distribution of task difficulty rating regarding the decision-making task for the two communication methods.

9. Personal Declaration

I hereby declare that the submitted Thesis is the result of my own, independent work. All external sources are explicitly acknowledged in the Thesis.

Date, Place:

7.6.22, Baar

Signature:

J. Bracher

Jana Bracher