The Influence of a Diverging Color Scheme on the Representation of Rock Glacier Dynamics in 3D Animations

GEO 511 Master's Thesis

Author
Flo Tiffany Flückiger
17-700-857

Supervised by
Dr. Tumasch Reichenbacher
Prof. Dr. Sara Irina Fabrikant
Prof. Dr. Markus Stoffel (markus.stoffel@unige.ch)
Alberto Muñoz-Torrero Manchado (alberto.munoztorreromanchado@unige.ch)

Faculty representative
Prof. Dr. Sara Irina Fabrikant

15.09.2022
Department of Geography, University of Zurich

Acknowledgments

This thesis would not have been possible without the help and support of a variety of people. Most importantly I would like to thank my supervisor **Dr. Tumasch Reichenbacher**, who guided me through this master's thesis and always provided valuable feedback in the many meetings we had and helped me explore my options when things did not go as planned. Further, I would like to express my gratitude to:

- **Prof. Dr. Sara Irina Fabrikant** who, together with Tumasch, helped me formulate my research question and hypotheses and provided valuable feedback for my concept talk.
- **Alberto Muñoz**, from the University of Geneva, who helped me with the many technical problems I encountered and with the procurement of the data. He always made time to help me when I got too frustrated and without him, I would have been completely lost during the photogrammetry process.
- **Prof. Dr. Markus Stoffel**, also from the University of Geneva. Even though we only stayed in contact during the beginning of my thesis, he was always very supportive. I also thank him for introducing me to Alberto and his work, which provided the basis for this thesis.
- **Dr. Ian Tanner Ruginski** and **Dr. Eun-Kyeong Kim** who provided me with valuable feedback for my thesis concept.
- the **Participants** of the survey. Without them the whole experiment would not have been possible, and I am very thankful for the 86 people who were willing to fill out my survey and thereby, be a part of my thesis.
- **Friends and Family** for proofreading this thesis and many other projects and most importantly for the moral support throughout my studies.

Abstract

Climate change has a large impact on a variety of fields and among the most heavily affected is the high mountain cryosphere, which includes rock glaciers. While it is a common approach for scientists to use image sequences such as 3D animations to detect changes in rock glacier dynamics, they are not always well perceivable for people with little knowledge about rock glaciers. However, with climate change being a topical subject, more and more people with little scientific background start to inform themselves on this issue. It is, therefore, important to provide information sources that are easily understandable by everyone. Existing literature suggests that a diverging color scheme could enhance the perception of rock glacier dynamics in terms of accuracy, perception time, and confidence. In this thesis, the influence of the diverging color scheme on these three variables is investigated by performing an online experiment in the form of a survey. The results show that the accuracy of the rock glacier perception was partly enhanced by the diverging color scheme. Further, it became evident that the addition of a diverging color scheme to a 3D animation led to significantly longer answering times as well as significantly stronger answer confidence. The results suggest a diverging color scheme partly enhances the perception of rock glacier dynamics in 3D animations, but more research is needed to further improve the visualization of rock glacier dynamics to make them easily perceptible for anyone with an interest in the topic.

Keywords: rock glaciers, diverging color scheme, 3D animations, accuracy, time, confidence, geovisualization

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Motivation

The climate change crisis is one of the greatest challenges the world is currently facing (Lange & Rand, 2021). While it has a large impact on a variety of fields, the high mountain cryosphere, with respect to glaciers, snow, and permafrost, is among the most heavily affected (Dimri et al., 2021). Consequences of climate change for the high mountain cryosphere may, among other things, be the release of organic carbon into the atmosphere from melting permafrost or paraglacial risks from melting glaciers (Mercier, 2021). Understanding the impact of climate change on the mountain cryosphere can be difficult and rock glaciers play a key role in understanding the extent of this impact (Ulrich et al., 2021).

Rock glaciers (Figure 1.1) are a specific type of permafrost, which is defined as ground that has been frozen (temperature at or below 0°C) for at least two consecutive years (Biskaborn et al., 2019). They are characterized by their slow high-viscosity flow-like downslope creep under deformation that may remind of regular glaciers (Strozzi et al., 2020). The annual movement rates of rock glaciers are rather little and vary from just a few centimeters up to a few meters (Steinemann et al., 2020). Changes in the movements of rock glaciers can be an indicator of environmental change (Strozzi et al., 2020). Therefore, it is interesting to observe rock glacier dynamics and draw conclusions about the impact of climate change on the mountain cryosphere in general.

The most common approach to observe rock glacier dynamics used to be with field observations. However, only using field observations to determine these dynamics is usually not sufficient due to the complex topography in rock glacier areas (Strozzi et al., 2020). A better option to monitor the movement of rock glaciers is with aerial images. Sequential aerial images have been used to detect change over time in various fields and even display the change in 3D through the application of photogrammetry (Feurer & Vinatier, 2018). This methodology has shown to be successful in detecting changes in rock glacier dynamics in a way that would not be possible in a two-dimensional setting or with field observations (Zahs et al., 2019).

Figure 1.1. A rock glacier on the Macun Plateau in the Swiss National Park (Derungs & Tischhauser, 2017).

While the analysis of sequential aerial images is becoming a more often used approach for observing the dynamics of regular glaciers through time (Paul, 2015), applying the same procedure to rock glaciers is not that common yet. This makes it attractive to employ similar methods in a yet to be investigated field. Observing rock glaciers over the last few decades may be particularly interesting as climate change is said to have accelerated rock glacier flows and changes in mass balance and thereby the thickness of the rock glaciers (Cusicanqui et al., 2021). Bringing the sequential aerial images together in an animation might also benefit the understanding of rock glacier dynamics as animations have shown to be able to reveal patterns that are not visible when static images are compared (Tyner, 2010).

With climate change being a topical issue, and hence many people being interested in it, it is important to create a form of visualization that is not only easily perceivable by scientists, but by everyone who has an interest in the topic (Diamond et al., 2020). However, it is a challenge to find a way to visualize rock glacier dynamics in a way that they are easily perceptible. As rock glaciers only move slowly over time (Steinemann et al., 2020), visually detecting changes can be difficult.

Fortunately, visualization science provides a variety of tools that may help to better visualize the rock glacier dynamics. For example, the visual variables (Bertin, 1983) combined with the standard map components can highlight associations between different variables (Opach et al., 2011). Being a powerful cartography tool, color, which includes multiple visual variables, can serve various purposes in visualization (Roth, 2017). One of which is highlighting change. In some cases, researchers want to highlight a specific feature or characteristic of a visualization. Highlighting becomes important when a simple sequence of images (static or animated) is not enough to visualize the desired aspects of changeability and there is a need for emphasis (Opach et al., 2011). It would, therefore, be interesting to see whether color can truly enhance the perception of rock glacier dynamics.

A promising way to visualize rock glacier dynamics is by applying a diverging color scheme. In most cases where rock glacier dynamics were visualized with color, a diverging color scheme was used to show the direction of change (increase/decrease) with different color hues and the magnitude of change with a stretch of color values (e.g., Abermann et al., 2010; Zahs et al., 2019). However, as of today, there are no studies investigating whether the application of a diverging color scheme truly enhances the perceivability of rock glacier dynamics in visualizations or not.

1.2 Aim and Research Question

The aim of this thesis is to find out whether a diverging color scheme can enhance the perception of rock glacier dynamics in a three-dimensional visualization. To investigate this, two 3D animations of a rock glacier on the Macun Plateau in the Swiss National Park are created from two-dimensional aerial images using photogrammetry. One of those animations has no addition of color and just shows the three-dimensional shaded relief, that was derived from the aerial images and the ensuing DEM (Digital Elevation Model), in black and white. The second animation has the same base layer, but the shape of the rock glacier is colored with a diverging color scheme indicating the range of increase or decrease of the rock glacier volume compared to the extent of the volume of the first aerial image of the sequence.

To assess whether the rock glacier dynamics are more easily perceptible with the addition of color, an experiment was designed, which was in the form of an online survey. As it is important that the visualization is comprehensible for anyone with an interest in the topic and not only scientists, people with little to no pre-knowledge of rock glaciers are included as participants. The participants were divided into two groups: Group 1 watched the animation with color and Group 2 the animation without color. Both groups then had to answer the same questions with the knowledge they gained from the animation.

This experimental setup allows to assess whether one form of visualization achieves better results than the other one. The goal thereby was to answer the following research question:

How does the application of a diverging color scheme on a three-dimensional animation of a rock glacier influence the perception of the rock glacier dynamics?

Following this research question, three hypotheses were formulated:

Hypothesis 1 (H1): The diverging color scheme helps to *better* understand the rock glacier dynamics.

Hypothesis 2 (H2): The diverging color scheme helps to *faster* understand the rock glacier dynamics.

Hypothesis 3 (H3): The diverging color scheme raises the *confidence* with which participants understand the rock glacier dynamics.

The hypotheses focus on the accuracy of the answers in the survey, the time participants needed to answer a question regarding the rock glacier dynamics as well as the participants' confidence with their answer. All three hypotheses favor the animation with the diverging color scheme. For the accuracy (Hypothesis 1), the hypothesis assumes that the diverging color scheme enhances the understanding of rock glacier dynamics, which is suggested by the fact that it is a common approach to use this scheme to visualize dynamics such as increase and decrease in volume (e.g., Immerzeel et al., 2014). Furthermore, in previous experiments, the diverging color scheme has proven itself to be more effective in terms of accuracy than a single-hue (e.g., black and white) scheme (Reda & Szafir, 2021).

Hypothesis 2 articulates that it takes less time to perceive the dynamics of the rock glacier when there is a color scheme indicating these dynamics. Previous studies have suggested that the addition of color on ortho-imagery may enhance perception including perception time (e.g., Hoarau & Christophe, 2017). The hypothesis is contradictory to the speed-accuracy trade-off, which states that the faster a question is answered, the less accurate the answer is, and vice versa (Heitz, 2014). Hence, if Hypothesis 1 predicts more accurate answers, the second hypothesis should predict longer answering times if this trade-off was considered. However, it may be possible that the black and white version is too complex to easily perceive the rock glacier dynamics, especially for people with little knowledge of rock glaciers. They would have to watch the animation more times to understand the dynamics than the participants that are in the group with the diverging color scheme, which would lead to longer answering times. Additionally, this trade-off usually applies to participants from a single group and not when comparing different groups (Heitz, 2014). Therefore, the second hypothesis is formulated as above.

The third hypothesis is about the participants' confidence. It is hypothesized that the animation with the diverging color scheme leads to a higher confidence in the participants' answers as they find it easier to recognize the rock glacier dynamics when the type and magnitude of change are indicated with color. This is because it is suggested that the confidence is stronger when there is more evidence supporting an answer (Koriat et al., 1980). With the addition of color, the participants do not only have the shaded relief to base their evidence on but have a further information source which should lead to a stronger confidence.

In summary, this thesis aims at investigating the influence of a diverging color scheme in 3D rock glacier animations on the perception of rock glacier dynamics. To do so, an online experiment helps to evaluate whether a diverging color scheme enhances the accuracy and time of answers to questions about the rock glacier dynamics as well as the participants' confidence with their answer.

1.3 Structure

This introduction section is followed by a chapter on the state of research on the visualization of rock glaciers and other geomorphological processes and phenomena. The focus lies on visualization in general as well as possible visualization variations, elaborating what the advantages and disadvantages of 3D over 2D and animated over static visualizations are. In a third section of the literature review, color in geovisualization is picked up and it is explained why the choice of the right color map is important and why the diverging color scheme is promising for the visualization of rock glacier dynamics. The third chapter elaborates on the data and methods that were used to create the animations as well as the experiment that was conducted to investigate the research question and hypotheses. This chapter is then followed by a section describing the results of the experiment. After that, these results are discussed and interpreted. Further, they are linked to the research question and the three hypotheses as well as to previous studies. Additionally, the limitations of the approach are explained. The last chapter, the conclusion, summarizes the findings of the thesis and gives an outlook on possible future work in the field of rock glacier visualization.

2 Related Work

A lot of research has already been done in the field of visualizing geomorphological processes as well as the application of color in geovisualization. This chapter elaborates on this research. In the first part, visualization of geomorphological processes, in general, is explained with a focus on visualizing landscape evolution, of which geomorphological processes, including rock glaciers, are a big part of. In the same section, current approaches to visualizing spatio-temporal phenomena are described. In the second section of this chapter, it is explained what variations exist regarding the visualization of geomorphological processes. The emphasis lies on the differences between 2D and 3D visualization as well as static and animated visualization. The respective approaches are compared, and the advantages and disadvantages related to the visualization of rock glaciers are listed. The third section focuses on color in geovisualization. There are various applications of color in geovisualization, but the focus of this section lies on color as a visual variable in general, numerical encoding, the creation of successful color maps, and, most importantly, the diverging color scheme and its application in visualization of geomorphological processes. Throughout the whole chapter, examples of studies that researched these topics are named and some are further elaborated on.

2.1 Visualization of Geomorphological Processes

Scientific visualization can be considered a very powerful tool to manipulate, represent and explore data and thereby gain an understanding of it (Silva et al., 2011). Visualization should help to effectively break down complex scientific information into something more easily understandable (Mitasova et al., 2012). This is becoming more important nowadays as the accessibility and quantity of data are steadily increasing and consequently, effective ways are required for data communication and analysis (Kelleher & Wagener, 2011). Hence, every scientist should have the intention to visualize their content in a way that is as easily understandable and as accurate as possible, so that even people that are not familiar with their research have no trouble understanding the matter (Crameri et al., 2020). In other words, visualization allows to simply demonstrate important information whereas describing the same information with words would be too complex to understand for people that are unfamiliar with the topic (Feng et al., 2019). Therefore, visualization is, in some settings, considered the most important component of research presentation and communication (Kelleher & Wagener, 2011).

Visualizing spatio-temporal processes can be especially difficult, and it has posed a challenging task for the cartographic community for some time now (Wilkening et al., 2019). As the name already implies, the temporal and spatial extent of a phenomenon are of interest. A spatio-temporal visualization, therefore, shows the sequence of states or events of a spatially spread phenomenon. The focus of spatio-temporal visualization lies on a variety of geographic phenomena and is generally required to be very accurate in order to investigate these phenomena (Opach et al., 2011). It is especially important as the visualization of spatio-temporal processes can not only help bring a subject closer to people that are not familiar with it but also help scientists to discover new information about processes (Wilkening et al., 2019). The applications of spatio-temporal visualization can be in an interactive setting as well as simple cartographic communication (Opach et al., 2011).

2.1.1 Visualizing Landscape Evolution

Landscape evolution is a scientific subject that incorporates how various geomorphological processes and phenomena, including rock glaciers, shaped various landscapes over time (Tucker & Hancock, 2010). Hence, geomorphological mapping has over time become a recognized approach for the reconstruction and documentation of landscape evolution (Seijmonsbergen, 2013). While geomorphological maps are applied in a variety of disciplines, investigating past states of a landscape can also help to make predictions about the future, making the mapping of geomorphological processes a very important tool in a variety of fields such as the projection of climate change scenarios (Seijmonsbergen, 2013).

To understand and study the geomorphological processes that shape landscapes and create landforms, it often helps to visually perceive them. Consequently, it is of great importance to create detailed and accurate visualizations (Smith et al., 2013). Tateoisan et al. (2014), therefore, list several principles that a visualization should follow in order to make geomorphological processes and phenomena more easily understandable. One of the three main points they mention is that a visualization should effectively use visual texture features like size, shape, and color as they are quickly perceptible. Second, data aspects of particular interest should be highlighted in some way. And last, there should be the possibility to interactively explore the data.

2.1.2 Current Approaches

With the advent of GIS (Geographic Information Systems), the visualization of spatio-temporal phenomena entered a new stage (Seijmonsbergen, 2013). GIS made it possible to visualize information based on real-world data in a photorealistic way and, if wished for, also in real-time (Wissen et al., 2008). This can be particularly effective for landscape evolution visualization, especially in combination with new technologies for mapping which revolutionized landform analysis with a greater level of detail and spatial extent (Mitasova et al., 2012). Another reason why GIS helped the visualization of geomorphological processes is that the coordinates of a map can be retained, which allows overlaying of different layers, e.g., DEMs of different years so that they overlay exactly (Smith et al., 2013).

A common approach to visualize geomorphological processes is with the comparison of DEMs such as DTMs (Digital Terrain Models) or DSMs (Digital Surface Models). Comparing DEMs of the same spatial extent but at different points in time allows to draw inferences on geomorphological processes which then can be visualized (Evangelidis et al., 2018). Additionally, landscapes can be modeled in 3D with the information a DEM provides, which is elaborated in section 2.2. With DEMs as a basis, a landscape can be visualized in various ways, for example, by creating a shaded relief, coloring according to slope or aspect, coloring according to landscape elements such as rock, glacier, or vegetation, and many more variations (Seijmonsbergen, 2013). Shaded reliefs have shown themselves to be a good way to visualize glacier dynamics (Wheate, 2012). Figure 2.1 shows an example of a scene once visualized as a shaded relief and once colored according to aspect. The scene is only a static image and, therefore, does not show a geomorphological process. To visualize the process a series of images of the same scene at different points in time would be necessary (Tucker & Hancock, 2010).

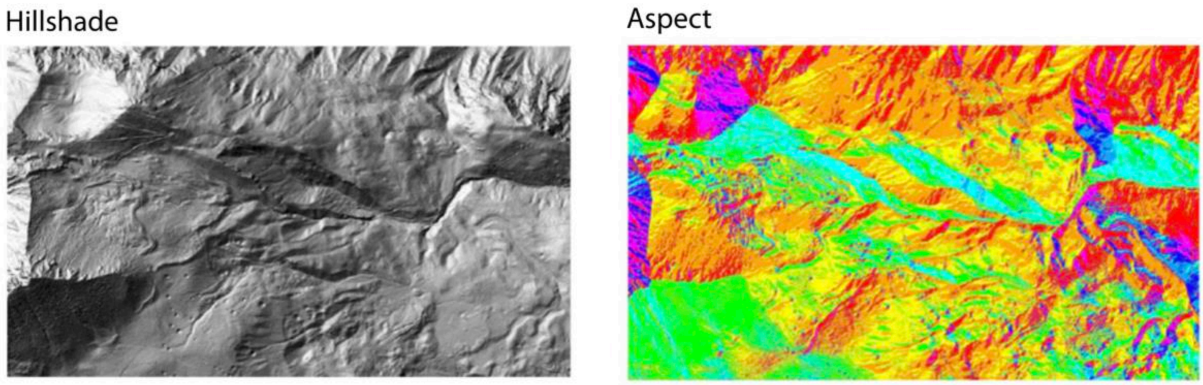


Figure 2.1. Shaded relief (hillshade) and aspect visualization of a scene. Adapted from Seijmonsbergen (2013).

Using DEMs, however, is only suitable for the analysis and visualization of processes that happen fast and not over the time span of several hundreds or thousands of years as DEMs only became available very recently compared to the timespan some geomorphological processes take (Werbrouck et al., 2011).

Nevertheless, some ways allow the creation of DEMs from other sources. With the use of photogrammetry, DEMs can be created from aerial images by extracting elevation information. The general photogrammetry approach uses at least two images with a fixed camera which then allows extracting the elevation information by taking advantage of the parallax effect (Hochschild et al., 2020). A similar approach is SfM (structure from motion), which is derived from photogrammetry, but the camera does not have to be fixed as long as the overlapping area between two pictures is large enough (Hochschild et al., 2020; Ullman, 1979). Because SfM offers a quite simple way to create DEMs through photogrammetry and DEMs have shown to be an effective way to visualize landscape evolution and the geomorphological processes associated with it, its main application is in geomorphological research (Langhammer & Vackova, 2018). For example, Immerzeel et al. (2014) researched glacier dynamics in the Himalayas and applied SfM to, among other things, visualize the mass loss of a glacier over time by calculating the DEM difference (Figure 2.2).

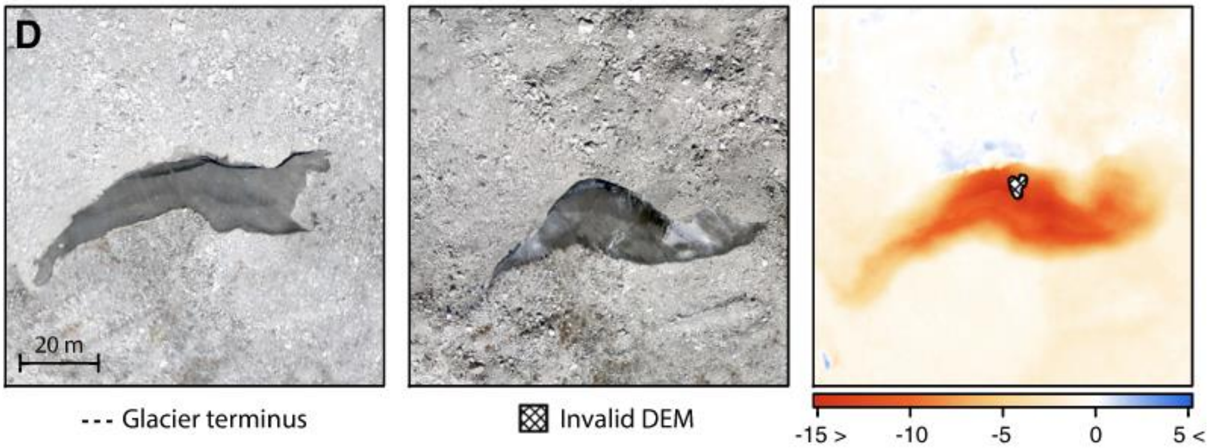


Figure 2.2. Visualization of a glacier in the Himalayas. Left: May 2013; Middle: October 2013; Right: DEM difference. Adapted from Immerzeel et al. (2014).

2.2 Visualization Variations

There are some variations in how geomorphological processes can be visualized. One of the two main differences that are often being discussed in current research is whether it is more useful to visualize them two- or three-dimensionally (e.g., Dübel et al., 2014; Philips et al., 2015; Wissen et al., 2008). The second point of contention is whether it makes more sense to visualize geomorphological processes with static images or animations (e.g., Ali & Motala, 2018; Opach et al., 2011). This section focuses on the advantages and disadvantages of these visualization variations in current literature.

2.2.1 2D vs. 3D

For a long time, 2D visualizations were the norm. However, 3D visualization has become more important with evolving technology (Philips et al., 2015). The opinions on whether two- or three-dimensional visualization is more effective are strongly divided. Several authors argue that there is little to no advantage of 3D over 2D. For example, Savage et al. (2004) conducted a study in which they investigated whether there was an advantage of 3D visualization over 2D visualization by asking participants to solve spatial tasks. Their results show no significant advantage of 3D visualization over 2D visualization. The same conclusion was also drawn by Dall'Acqua et al. (2013). In their experiment of visualizing permafrost, the outcome even showed 2D visualization to be more effective than 3D visualization.

Nevertheless, there are studies that show different results, especially concerning geomorphological processes. Dübel et al. (2014) say that whether to use 2D or 3D visualization depends on the application the visualization is used for. In geomorphological visualization, 3D visualization has often shown to be more effective than 2D visualization. For example, Mitasova et al. (2012) claim that, especially in geomorphology, 3D visualization is more valuable than 2D visualization because 3D patterns and relationships are crucial for understanding phenomena and depicting landscape features. Wissen et al. (2008) also underline the importance of 3D visualization in geomorphology as it may contribute to an advanced understanding of geomorphological processes. Another reason why they argue pro 3D visualization and against 2D visualization is that 2D visualizations often abstract reality considerably and therefore have a lower level of realism (Wissen et al., 2008). James & Robson (2012) support the arguments of the authors of the last two studies mentioned by saying that 3D visualization in recent years has significantly increased the understanding of a variety of geomorphological processes.

As one example of a 3D visualization in a geomorphological field, Zahs et al. (2019) conducted a three-dimensional change analysis of a rock glacier. They find that by capturing the data in 3D, they were able to quantify surface change as well as its direction, which would not always be possible in a 2D setting. Furthermore, Zahs et al. (2019) propose that volumetric calculations based on the mentioned surface changes could provide further insights into rock glacier dynamics.

2.2.2 Static vs. Animated

Besides the discussion of whether two- or three-dimensional visualization is more effective for the visualization of geomorphological processes, there are also inconclusive opinions on the usage of either static images or an animation, which displays change over time dynamically. The main difference between static and animated maps is that static maps display all of their information at once, whereas animated maps display information over time (Hochschild et al., 2020). This topic has especially become important since paper maps started to get replaced by digital platforms like GIS (Seijmonsbergen, 2013). The main reason why animations became more important is that they sometimes reveal patterns that are not visible on static maps. However, before creating an animation instead of a static map one should always ask oneself whether there is true added value behind it or whether the animation is only more eye-catching than a row of static images (Tyner, 2010). The success of an animation is determined by several elements related to the animation's design (e.g., visual variables) as well as the characteristics of the data being represented such as complexity and resolution (Ali & Motala, 2018).

There are several studies that make direct comparisons between static and animated maps. One of these studies was conducted by Dawood & Motala (2015). In their experiment, they asked participants to answer questions after they had either viewed an animated or static map. Their results show that there are advantages and disadvantages to static maps over animated ones and vice versa but neither is truly superior. An experiment conducted by Ali & Motala (2018) obtained the same inconclusive results with either map being stronger than the other one in certain aspects. One thing they note, which is important for this study, is that the animation generally performed better in depicting spatial change. It is therefore not surprising that animations have become a crucial tool for the visualization and analysis of geomorphological processes (Mitasova et al., 2012).

Rastner et al. (2016) created a 3D animation of a glacier in Switzerland. For their study, they did not even consider using static images for visualization purposes. They made it clear from the beginning that one of their goals was to visualize the glacier dynamics over time with an animation. They found that the animation demonstrates the glacier dynamics very well, making them easily comprehensible to the wider public.

2.3 Color in Geovisualization

One of the most powerful tools in visualization is color. It has the power to transform information into meaning (Crameri et al., 2020) and is at the same time both an art and a science (Samsel et al., 2018). For humans, color vision is a rapid way to acquire information (Zhou & Hansen, 2016). It is therefore especially important to choose wisely which colors are used and for what purpose. Science has become increasingly widespread, not only in the scientific community but also in the wider public, making the choice of good color maps even more essential to communicate scientific purposes (Crameri et al., 2020). For example, hazard maps should be quickly perceptible to scientists as well as society. With a bad choice of color, this may be tricky. This section focuses on color in geovisualization and more precisely on its application as a visual variable, how numeric data can be encoded with color, the characteristics a successful color map should have, and lastly, the diverging color scheme and its advantages over other color schemes, especially in the field of 3D visualization of geomorphological processes, is introduced.

2.3.1 Color as a Visual Variable

The visual variables were introduced by the French Cartographer Jacques Bertin (1983). He initially described seven visual variables: location, size, color hue, color value, texture, orientation, and shape. These visual variables are conventional means to enhance cartographic design and are essential in geovisualization (Roth, 2017). Color is represented in two of these seven visual variables: color hue and color value. That is because, with color, different aspects can be emphasized. Color hue refers to the color variation, for example, red and blue are two different color hues. Color value describes the lightness of a color being used, for example, light blue and dark blue (Smith et al., 2013). Figure 2.3 illustrates the two visual variables in an easily comprehensible way.

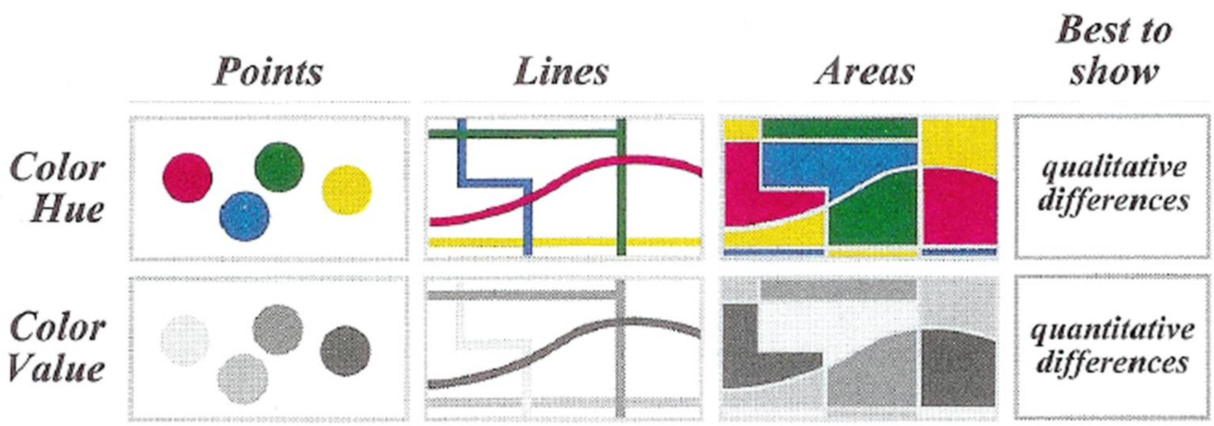


Figure 2.3. The visual variables color hue and color value. Adapted from Krygier & Wood (2005).

From Figure 2.3, it can be seen that color hue is often used to depict qualitative differences, while color value is more commonly used to describe quantitative differences. For color hue, this means that its main application lies in displaying different categories in a visualization. For example, Hoarau (2011) used different color hues to distinguish between map elements. She used blue hues for water bodies, green for vegetation, and purple for buildings. Color value is more suitable to describe quantitative differences like ordinal or numerical information (Roth, 2017). Brewer et al. (1997), for example, used color value to map mortality rates, among other things, in shades of grey. A dark grey area implied a higher mortality rate than a light grey one. For maps like that, color value is an appropriate and intuitive visualization tool. In conclusion, the visual variables of color are very powerful visualization tools that allow mapping qualitative as well as quantitative differences in data.

2.3.2 Numerical Encoding

Color has many purposes in visualization. An important one is the mapping of numbers to colors, in other words, encoding numerical data (Moreland, 2009). Examples are maps that show different temperatures or changes in height in different colors in a way that they are easily perceptible by users (Wiesmann et al., 2009). As already explained in the previous section, using color value is a better way to do so than color hue as it is more suitable to visualize quantitative data. A combination of both is possible as well. Nevertheless, only using color hue as an encoding should be avoided for numerical data as it is not very intuitive (Silva et al., 2011).

There are a few rules that should be followed when encoding numerical data with color, the most important one being perceptual uniformity. This means that the differences in the underlying numerical data should be perceived as equally different in the encoded color values (Szafir, 2018). Otherwise, the visualization would not be effective. This is especially relevant for continuous color maps to avoid misperceptions (Moreland, 2009). The importance of this is also underlined by several other authors. For example, Silva et al. (2011) emphasize multiple times how important it is for a user's perception that equal steps in the data also correspond to equal steps in the color map. It is not only important to encode the numerical values with equal differences in the color map, but also to make sure that these differences can be well perceived.

Therefore, it is central to maximize the number of noticeable differences along the scale (Levkowitz & Herman, 1992).

Another point that should be paid attention to when encoding numerical data is that the order of the chosen colors should be perceived as having the same order as the underlying data (Silva et al., 2011). Maintaining a logic order is also essential when encoding numerical data not only with color value but in combination with color hue. The order of the different color hues should make sense and be easily rememberable, for instance going from light to darker values (e.g., yellow-orange-brown) (Smart et al., 2020), or, taking the color encoding of temperature values as an example, it is intuitive to use "cold" and "warm" colors to create a scale representing cold and warm temperature values (Levkowitz & Herman, 1992).

2.3.3 Color maps

The most important part of encoding numerical values, as well as using color in geovisualization in general, is choosing a suitable color map, which is often also referred to as color scheme or color palette (Cheng et al., 2019). As explained above, color maps are so important because they have the ability to improve the effectiveness and efficiency of data perception (Zhou & Hansen, 2016). However, when choosing an appropriate color map, it is not just about aesthetics and creating a visually attractive map but also making sure that the added color adds further insight (Silva et al., 2011). An ill-designed color map may lead to confusion and substantially impair the effectiveness of a message (MacDonald, 1999; Zhou & Hansen, 2016). Hence, color maps are often most effective when the focus lies on beauty in terms of clarity and functionality, meaning that a lot is communicated with little (Grainger et al., 2016).

Choosing the right color map is difficult because not all color maps are suitable for all science domains and visualization types (Samsel et al., 2018; Smart et al., 2020). Nevertheless, choosing color maps has become passive for most scientists, and not the same kind of attention is paid to it as to other data methodologies even though it is just as important (Crameri et al., 2020). The choice is often based on intuition instead of perceptual principles (Grainger et al., 2016). However, this is not a smart approach because many aspects need to be considered when choosing a color map, that are often forgotten which sometimes can have crucial consequences.

The most commonly chosen color map in scientific visualization is probably the rainbow color map (Figure 2.4), which includes most of the saturated colors on the visible spectrum (Moreland, 2009). However, experts see a lot of issues with the rainbow color map. These problems can be divided into the following three categories: sensitivities to color deficiency, unnatural ordering, and irregular perception (Moreland, 2016). These things are essential for all color maps and are now further explained by demonstrating them with the rainbow color map.

Figure 2.4. The rainbow color map (Moreland, 2016).

One thing that is often forgotten is that many people are color blind or have color vision deficiencies (CVDs). It is estimated that 1 in 12 men and 1 in 200 women (8% and 0.5% of the world's population, respectively) are affected by CVDs (Alam et al., 2022). This corresponds to approximately 300 million people worldwide living with CVDs. Scientific results should be accessible to everyone that wants to access them. When a color map cannot be accurately perceived by a part of the population this is not the case. Hence, it is important to create CVD-friendly color maps in order not to exclude anyone from that opportunity (Crameri et al., 2020). For example, the rainbow color map is particularly CVD-unfriendly. Figure 2.5 illustrates how people with the three most common CVDs perceive the rainbow color map compared to how they perceive a color map that is CVD-friendly. Looking at the figure, it becomes evident that the rainbow color map does not allow people with CVDs to extract all the information that may be visualized.

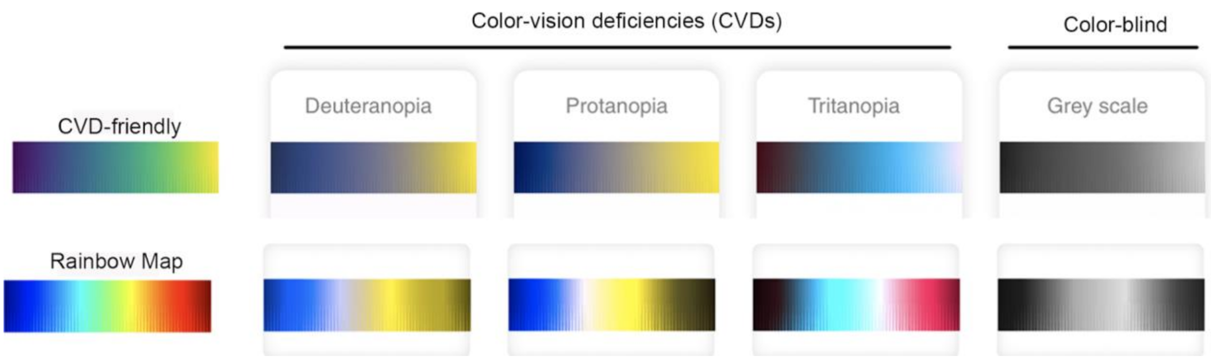


Figure 2.5. A CVD-friendly color map and the rainbow color map displayed how people with the three most common as well as completely color-blind people perceive them. Adapted from Crameri et al. (2020).

The second mentioned problem of the rainbow map is the unnatural ordering which was already touched upon in the numerical encoding section. It is crucial to have an intuitive order in a color scheme. The rainbow map does not have that. The standard RGB (Red-Green-Blue) colors are dominating and may divert from the underlying message (Crameri et al., 2020). Additionally, yellow, the brightest color, is the most eye-catching (Wolfe & Horowitz, 2017). However, it is neither at the center nor at one of the ends of the color map, which may lead to highlighting less important features more strongly (Crameri et al., 2020).

Irregular perception is listed by Moreland (2016) as the third issue with rainbow color maps. This means that the color hue changes in the map are perceived to be of different lengths, e.g., the yellow zone appears to be much shorter than the green one (Borland & Taylor II, 2007). This does not correspond to what was explained in chapter 2.3.2, where it was stated clearly that equal steps in data differences should be represented by equal steps in the color ramp.

For 3D visualizations, there is another issue that becomes important when finding a suitable color map. In 3D visualizations that are combined with color, problems can arise with the shading that may occur in a three-dimensional setting (Silva et al., 2011). This is the case when the color map depends strongly on the brightness of the colors. Hence, a color may be perceived as darker than it actually is due to the shading of the 3D surface, which is especially problematic in greyscale maps (Smart et al., 2020). The best way to bypass this problem is by using a color scheme with mainly bright colors and avoiding dark colors.

Many problems may arise when choosing or creating an appropriate color map. Moreland (2009) created guidelines for an ideal color map design for scientific visualizations. The guidelines give a good summary of what has been discussed in this section. In his guidelines he lists the following six criteria:

- The color map being used is aesthetically pleasing.
- The perceptual resolution of the color map is maximized.
- There is as little as possible interference with the shading of 3D surfaces.
- People with vision deficiencies have no problem seeing the whole spectrum of the color map.
- The order of the colors is intuitive for all people.
- Perceptual uniformity is ensured.

Hence, designing an effective color map is anything but a trivial task and requires a lot of expertise. Smart et al. (2020) go as far as recommending people with little experience not to try and design their own color map but instead use pre-constructed color schemes to prevent confusion and misperceptions.

There are several tools that assist in the selection of an appropriate color map. One of these tools is ColorBewer (colorbrewer2.org). ColorBrewer was developed by Cynthia Brewer and provides a variety of different color schemes. The user has the opportunity to choose the nature of the data (sequential, diverging, or qualitative color scheme, which is further explained in the next subsection) as well as the number of classes that should be represented in the color map. Additionally, the user has the option to choose a CVD-friendly color map. ColorBrewer then suggests a variety of suitable maps and the user can decide which one is the most aesthetically pleasing to them. ColorBrewer has proven itself to be a great tool for finding suitable color maps and is recommended by several authors, e.g., Crameri et al. (2020) or Smith et al. (2013).

2.3.4 The Diverging Color Scheme

In general, color maps can be divided into three categories: qualitative, sequential, and diverging. A qualitative color scheme uses different color hues to distinguish between data sets. Sequential color maps represent numerical values that have a logical order from a low to a high value such as velocity or material density. The diverging color scheme can be used to visualize values that have a midpoint, such as temperatures above and below 0°C (Kulesza et al., 2017). A diverging color scheme is actually the combination of two sequential color schemes into a single color map (Roth, 2017). As the focus of this thesis lies on the application of a diverging color scheme and on how it may improve the perception of rock glacier dynamics, this subsection elaborates on this specific type of color map.

To create a diverging color scheme the visual variables color hue and color value are combined. The color hue indicates the direction in which a value differs from the midpoint and the color value indicates how strongly a value deviates from the midpoint (Roth, 2017). The midpoint of a diverging color scheme is usually colored in an unsaturated color such as white or a light yellow (Moreland, 2009). Brewer et al. (1997) suggest four examples of diverging color schemes (Figure 2.6). All four color maps have the midpoint represented in white and the two endpoints in colors that are distinctly perceivable by everyone, including people with CVDs. The color schemes are presented in the form of a classified color scheme. However, the values could also be interpolated to fit on a continuous color map (e.g., Figure 2.7).

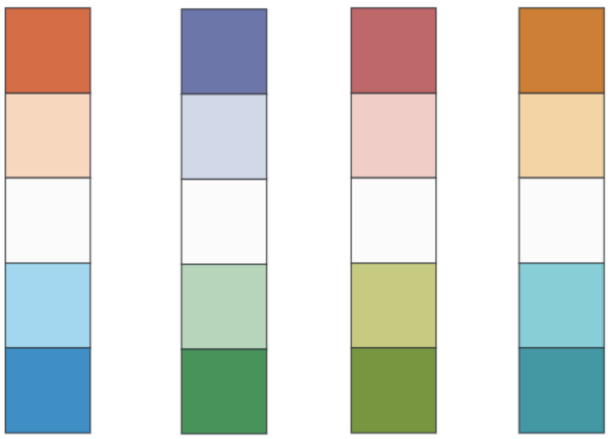


Figure 2.6. Examples of (classified) diverging color schemes (Brewer et al., 1997).

The most important part, besides having a clearly distinguishable midpoint, is that the color hues used for the two represented categories are contrasting (Grainger et al., 2016). The color map that is probably used most commonly for diverging color schemes is the one going from red to white to blue (Kulesza et al., 2017), which can be seen in Figure 2.7. There are several potential reasons for that. First, the red-white-blue scheme is CVD-friendly (Moreland, 2016). Additionally, some diverging color maps may lack natural ordering of colors if the two chosen color hues do not have "low" and "high" associations. This can be achieved by using "cool" and "warm" colors (Moreland, 2009). In general, red and yellow are considered warm colors, and blue and blue-greenish color hues are often identified as cool across different cultures (Hardin & Maffi, 1997). Hence, by using red and blue as the two color hues applied in a diverging color scheme, the order of the colors becomes more natural. Another reason why the red-white-blue color map is so popular is that the colors generally have a high luminance which makes the color scheme well suitable for three-dimensional visualizations as dark colors may interfere with the shading of 3D surfaces, as already explained above (Kulesza et al., 2017; Moreland, 2016).

Figure 2.7. The most commonly used diverging color scheme (red-white-blue). Adapted from Moreland (2009).

As the diverging color scheme, especially the red-white-blue one, has shown itself to be successful, it has unsurprisingly been wildly applied in various application fields including the visualization of geomorphological processes, such as rock glacier dynamics. One example was already shown above in section 2.1.2 (Figure 2.2). The color map chosen by Immerzeel et al. (2014) to illustrate the DEM differences of a glacier is a red-white-blue diverging color scheme. Another example where this type of color map was used to visualize changes in glacier dynamics is by Bhattacharya et al. (2021). They also illustrated differences between DEMs with the diverging color scheme. However, they used a light yellow as the color for the midpoint instead of white. An example of a visualization of specifically rock glacier dynamics with a diverging color scheme is by (Abermann et al., 2010). Again, they showed the DEM differences with this color map (Figure 2.8).

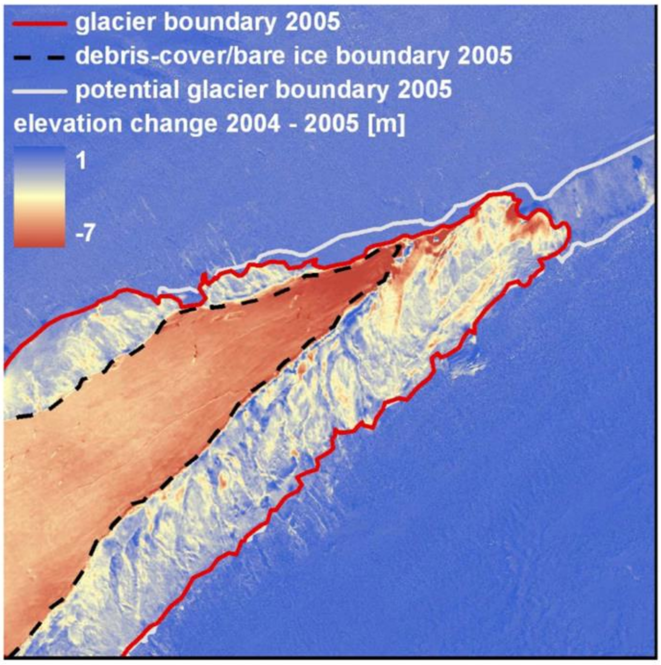


Figure 2.8. A diverging color scheme applied to illustrate DEM differences of two years of a rock glacier. Adapted from Abermann et al. (2010).

There are some studies that focus on the addition of color schemes on 3D models in construction (e.g., Chang et al., 2009; Chen et al., 2013). They focus on adding a fourth dimension to three-dimensional models with the addition of color. While in this thesis the goal is not to add a fourth dimension to a three-dimensional rock glacier model but rather highlight the change that would already be visible, some findings can be applied as well. For one, it was found that a diverging color scheme works well, especially for sequential data (Chang et al., 2007). Furthermore, in an experimental setup it was discovered that a diverging color map eases the understanding of what is represented in a way that participants need to use less mental effort to understand a model with a useful color scheme, which in the case of this study was a divergent one (Chang et al., 2009).

All in all, the diverging color scheme has shown itself to be effective and fulfills all the essential criteria a successful color map should have (Moreland, 2009). It has been applied multiple times in the field of visualizing geomorphological processes in the past years as well as in other fields. While some studies are comparing the success of various color schemes against each other (e.g., the rainbow color map to other color maps), as of now, there has little to no research been done on whether the dynamics of a rock glacier can be better represented by applying a diverging color scheme or with no color map at all, simply showing, for example, an orthophoto. However, by what is known, it is suggested that a diverging color scheme may improve the understanding of processes in general but also specifically applied to rock glacier dynamics. This thesis hopes to gain knowledge regarding this research gap.

3 Data and Methods

In this chapter, it is explained which data were used and which approach was applied to investigate the research question and the hypotheses. The focus first lies on the data that was used to create the 3D models and in the second part on the design of the 3D animations based on these data as well as the application of the diverging color scheme. In the third part of this chapter, the compilation and setup of the online experiment are described.

3.1 Data and Preprocessing

The main data being used for this thesis are aerial images from *swisstopo*, Switzerland's federal office of topography, of a rock glacier on the Macun Plateau in the Swiss National Park. This rock glacier was chosen as swisstopo offers aerial images with good coverage of the rock glacier over a large time span (from 1939 to 2003). From 1999 on, the images are available in color. The other images are only available in black and white. The resolution of the images varies between 13x13 cm and 23x23 cm. In other words, the images have a very high spatial resolution. Images are not available for every year between 1939-2003. In fact, the swisstopo archive only stores aerial images covering the Macun Plateau for twelve years: 1939, 1946, 1959, 1961, 1973, 1978, 1985, 1991, 1997, 1999, 2000, and 2003.

For this thesis, not only images from the rock glacier itself but also from the surrounding area were collected from swisstopo in order to enhance the photogrammetry procedure which was explained in section 2.1.2. This led to a total of 287 images being used to create 3D animations of the rock glacier on the Macun Plateau between 1939-2003. A complete list of all the images that were considered in this thesis can be found in Appendix A. Figure 3.1 shows an example of an aerial image that was used. In this figure, the outlines of the rock glacier on the Macun Plateau are well visible.

As can be seen in Figure 3.1, the aerial images all had a frame. Sometimes, there were notes on the frame or other additional information. This frame had to be cropped off as the photogrammetry software would have a problem with it. This was done with Adobe Photoshop (version 23.5.0).

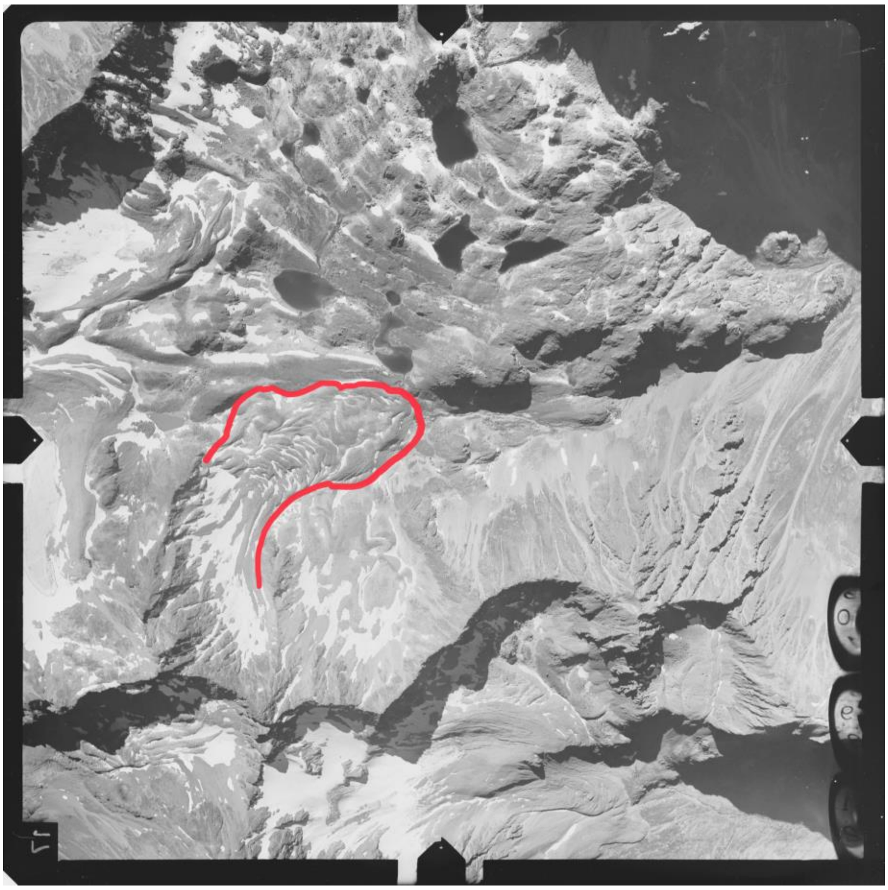


Figure 3.1. Aerial image from swisstopo of the rock glacier on the Macun Plateau in the Swiss National Park (1939). The outline of the rock glacier is marked in red. Image number: 19390740040091.

To combine the models of the individual years into a single animation, there is a need for ground control points (GCPs). GCPs are defined points on the surface of the Earth of a known location (Martínez-Carricondo et al., 2018). They are used to geo-reference the aerial images. Such GCPs exist for the rock glacier on the Macun Plateau. Originally, the idea was to work with those existing GCPs and an orthophoto that was created in the Bachelor's Thesis of Derungs & Tischhauser (2017) with a spatial resolution of 8 cm. However, the GCPs were not visible on the orthophoto. By the time this orthophoto could be accessed, it was already mid-October and the first snow on the Macun Plateau had already fallen. This made it not particularly useful to do fieldwork and generate new GCPs. Therefore, it was necessary to come up with a new idea.

Instead of working with the ground control points, only the orthophoto was used. On the orthophoto, distinguishable points were defined. These so-called *reference points* must be stable features around the moving rock glacier mass, for example, a stable boulder. To verify that the locations were indeed stable, the images of 1939 and 2003 were compared and it was made sure that the reference points are present in both years and are still in the same position. This was validated by comparing the relative distance of the reference points to another well-known stable position. In total, six reference points were defined around the rock glacier. The locations of the reference points can be seen in Figure 3.2. Once the reference points were defined, the coordinates of those points were extracted from the orthophoto, which is georeferenced.

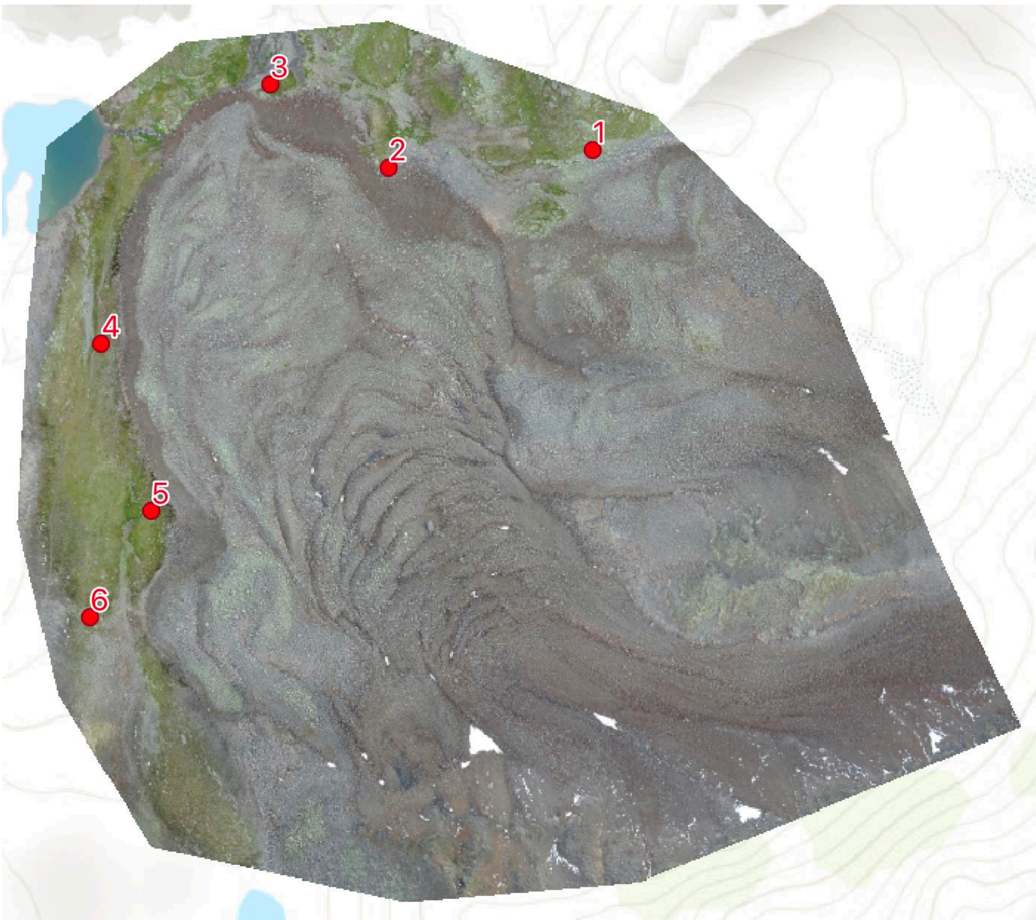


Figure 3.2. Locations of reference points around the rock glacier on the Macun Plateau on the orthophoto from Derungs & Tischhauser (2017).

As the goal of the thesis is to investigate the influence of a diverging color scheme on a *three-dimensional* animation, it was necessary to not only collect the 2D coordinates, but also the elevation as the third dimension. The data for the elevation was obtained from *swissALTI3D*. This is a digital elevation model (DEM) which represents the surface of Switzerland without development and vegetation very precisely with a resolution of 0.5 m (swisstopo, n.d.). The part of the DEM showing the Macun Plateau was downloaded and opened in ArcGIS Pro (version 2.9.0). Then, the six reference points were identified in the DEM by their coordinates. From that, it was possible to extract the elevation value from the DEM. The three-dimensional coordinates of all the reference points were then stored in an Excel sheet for later usage in the photogrammetry software.

In summary, three types of data were used: 287 aerial images provided by swisstopo, an orthophoto created for the Bachelor's Thesis of Derungs & Tischhauser (2017), and the DEM swissALTI3D also provided by swisstopo. A few preprocessing steps have been performed, namely cropping the border of the aerial images, defining reference points, and extracting their coordinates as well as the elevation of these points.

3.2 3D Animations

For the creation of the 3D animations themselves, several steps were necessary. First, the two-dimensional images were used to extract three-dimensional information about the rock glacier and its surrounding area by applying photogrammetry. These three-dimensional outputs were then further processed into the right format for the creation of an animation, namely shaded reliefs for the display without color and difference images for the display with color. In a third step, the prepared layers were combined into two 3D animations (one with color and one without).

3.2.1 Photogrammetry

The photogrammetry software being used in this thesis is Pix4Dmapper (version 4.6.4). It is a program that uses stagnant 2D images to generate point clouds as well as 3D models and orthophoto maps (Barbasiewicz et al., 2018). The input for every individual 3D model was the cropped images for the respective year. In the first step, a coordinate system was assigned. The coordinate system worked with in this thesis is WGS84 / UTM Zone 32 N (EPSG: 32632). The software tries to find tie points between the images. The minimum number of tie points to calibrate an image defined by the software is two. To obtain an a little more accurate result, the minimum number of tie points was set to three. This worked for most years. However, some years had a lot of uncalibrated images with this setting. For those years, the minimum number of tie points was reset to two. The years this had to be done were: 1946, 1985, 1991, 2000, and 2003.

After the aerial images were calibrated, a point cloud was created with the Pix4D software. This software gave a first glimpse of what the final 3D model may look like. The next step was to assign the coordinates of the reference points to the calibrated images. To do so, the rock glacier had to be found in the point cloud first. Then a point as close as possible to a reference point had to be clicked in the point cloud. Subsequently, all the calibrated aerial images this point can be found in showed up in a side panel. The exact reference point then had to be determined and clicked in each image and then the respective three-dimensional coordinates could be manually assigned. This had to be repeated for all the six reference points for every year. After that, the point cloud had to be created a second time, with the newly assigned coordinates.

Unfortunately, a few years already had to be discarded in this step: 1939, 1946, 1991, 1997, and 2000. There were two reasons for that. In the case of the years 1946 and 1997, the rock glacier was not visible at all in the point cloud. The reason for that is that the images covering the rock glacier were not calibrated as the software was not able to find enough tie points. In the point clouds of the other three years, the rock glacier was only partially visible. In the case of 1939, 1991, and 2000, half of the rock glacier and the locations of the reference points were cut off and generally, the rock glacier, as well as its surrounding area, were full of large holes, making it impossible to find the locations of the reference points.

After the point clouds were created again with the newly assigned coordinates, the point clouds had to be edited. Many points were not correctly positioned in the cloud (far too high or far too low) due to several reasons, for example, snow coverage or shadows. All these points had to be selected and disabled in Pix4D as they may have falsified the final model. Then the digital surface model (DSM) and orthophoto map were created in a final running step.

After this step, two more years had to be discarded. 1985 had to be disposed of because in the final orthophoto map the rock glacier is not visible anymore. Before the last step, it was already at the very edge of the point cloud and unfortunately, it is completely cut off in the final version. Therefore, there was no use for this year anymore.

Another year that had to be discarded after this last step is 1961. In the case of this year, the orthophoto map obtained (Figure 3.3) does not show the rock glacier truthfully. Although the whole extent of the rock glacier is visible, it appears that the upper part of it is somehow twisted and facing in the opposite direction of where it is facing in reality. Hence, it would not have made sense to include 1961 in the further process.

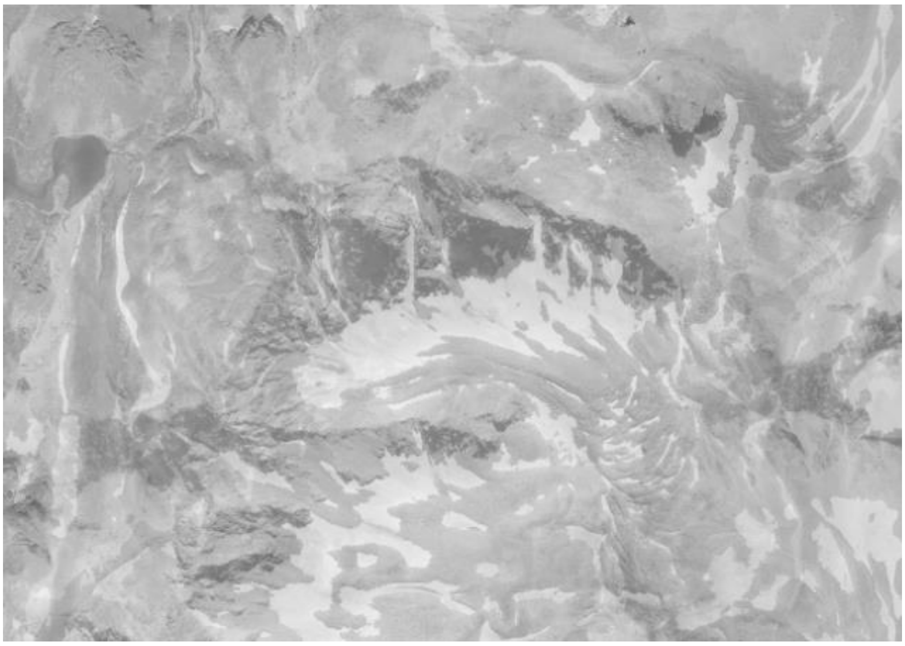


Figure 3.3. Orthophoto mosaic of the year 1961 created by Pix4D software. The representation of the rock glacier in this image is inaccurate as it is somehow split in half and twisted.

For the remaining years (1959, 1973, 1978, 1999, and 2003) the photogrammetry software worked well, and the results were able to be exported for further preparation for the animation, which is explained in detail in the next section of this chapter.

3.2.2 Preparation of Layers for Animations

To prepare the layers for the animations several steps were necessary. The main data used for this process were the DSMs created with Pix4D. They were all loaded into ArcGIS and firstly, cropped to a smaller area around the rock glacier to reduce processing time. To gain a first overview, shaded reliefs for all the years were created. It became visible that most of them were not smooth and there were quite a lot of bumps and holes in the shaded reliefs. Therefore, the first preparation step was to smooth the DSMs to try to get rid of those bumps and holes. To do so, a series of steps in ArcGIS Pro was required. For each year, the slope of the area around the rock glacier was calculated using the DSM of the respective year. The areas with a slope higher than 55° were extracted and a buffer of 1m was added. From the original DSM, these buffered areas were removed, and the DSM was re-calculated without those areas. A detailed description of all the exact functions used for the smoothening procedure can be looked up in Appendix B.

The size of the buffer has a big impact on how much the DSM will be smoothed. Shaded reliefs from DSMs with four sizes of buffers were created: 0.5m, 1m, 1.5m, and 2m. 2m generally smooths the surfaces the most while 0.5m smooths the least. However, the enhancement of the quality of the shaded reliefs depending on the buffer size differed between the years. For 1959 and 2003, there were only a few differences between all the buffer sizes. For 1973 and 1978, a 2m buffer obtained the best result while a 0.5m buffer was still a not very good improvement to the original DSM. For 1999, the 0.5m buffer obtained a very good result but the larger buffers led to a lot of inaccuracies and pixelation. A buffer of 1m, however, was still acceptable. These differences between the years made it not particularly easy to choose a buffer. Based on subjective judgment, it was decided that a buffer of 1m would obtain the best general result and therefore, a 1m buffer was applied in the smoothening procedure.

The resulting DSMs were then cropped to the area of the rock glacier as the animations will focus only on this specific area. To do so, a polygon with the rock glacier outline was manually drawn by using the orthophoto map of 2003 as a reference. For simplification reasons, only one polygon was drawn and not one for each year. As rock glaciers are moving constantly, they may also slightly change their shape (Frauenfelder & Kääb, 2000). Therefore, it would have been scientifically more accurate to draw a polygon for each year. However, this would have complicated the process of bringing the individual years together, and hence, it was decided to only use one polygon. This is justifiable because even though the shape of the rock glacier may have changed over the years, the change would still be very small (Steinemann et al., 2020).

Additionally, the cell sizes of the resulting DSMs in the shape of the rock glacier on the Macun Plateau were all set to 1m. There are two reasons for that. For one, having the same cell size for each DSM also simplifies the process of bringing the individual years together for the animations. Secondly, a resolution of 1m, which is lower than the original resolutions of just a few centimeters, reduces processing time but still preserves a lot of information.

Having a look at the minimum and maximum elevation of the rock glacier-shaped DSMs (Table 3.1), it becomes evident that while the minimum elevations all seem to be in a similar range (within ~ 5m), this is not the case for the maximum elevation. Whereas the maximum elevation for the years 1973 and later are also in a similar range (within ~ 4m), the year 1959 has a maximum elevation of 2945.57m, which is almost 50m higher than the second-highest maximum elevation (2895.91m, 1978). Such a drastic change is not possible over a short period of not even 20 years as rock glaciers are known to have very slow movement rates (Steinemann et al., 2020). Including the year 1959 as it was, would have led to a major distortion in the difference images as 1959 is the basis for these images (explained later in this section). Therefore, this had to somehow be corrected.

As a trend of the change in maximum elevation over time is not clearly distinguishable, it was not possible to therewith determine an approximate value of the actual maximum elevation in 1959. It was decided to use the mean value of the maximum elevations of the other four years as a plausible alternative maximum elevation for the year 1959. Mean values are generally sensitive to outliers (Leys et al., 2013). Therefore, they are not always a good method to try to estimate another value. As the four values considered for the mean are however all in a small range, it is justifiable to do so. The mean value calculated and used as an alternative maximum elevation value for the year 1959 is 2892.74m.

Table 3.1. Minimum and maximum elevation [m] in the rock glacier-shaped DSMs of all the years.

Year	Min. Elevation [m]	Max. Elevation [m]
1959	2585.32	2945.57
1973	2590.45	2891.86
1978	2590.04	2895.91
1999	2588.32	2891.6
2003	2589.56	2891.57

The upper part of the rock glacier, which is the part where all the extremely high values occur, was extracted, and rescaled with the following equation in the raster calculator function of ArcGIS Pro:

$$r = \frac{(g - g_{min}) \times (r_{max} - r_{min})}{g_{max} - g_{min}} + r_{min} \qquad (1)$$

In this equation (1), r is the rescaled raster and g is the original raster grid. min and max stand for the minimum and maximum value in either r or g, respectively. This formula is from the official Esri support website (Esri, 2016). As, in this particular case, the raster was intended to be rescaled only in one direction (lower upper limit), r_{min} was equal to g_{min}. For r_{max}, the with the mean function determined value of 2892.74m was used. The rescaled raster was then re-combined with the lower part of the rock glacier, resulting in an enhanced version of the rock glacier in the year 1959.

As the time gaps between years with data of the rock glacier are very irregular, it was decided to perform a temporal interpolation. For simplification reasons, this interpolation was conducted linearly even though this does not necessarily represent actual rock glacier behavior (Rastner et al., 2016). The goal was to have time gaps of approximately five years while making sure that all the original years were still included which sometimes led to time gaps of four or six years. Between 1959 and 1973 two layers were interpolated (1964 and 1969) and between 1978 and 1999 three layers (1983, 1988, and 1993). They were calculated by using the raster calculator function in ArcGIS Pro with the following equation:

$$x = a + \frac{b - a}{c} \times d \tag{2}$$

In this equation (2), x is the interpolated DSM, a is the DSM of the earlier year, b the DSM of the later year, and c is the number of years between the year of a and the year of b. d represents the number of years between a and x. The fraction represents the amount of difference between a and b per year which is then multiplied by d and added to a to get the linearly interpolated layer.

Having a DSM layer for approximately every fifth year between 1959 and 2003, the layers for the 3D animations were able to be prepared. For the animation without color, shaded reliefs were created. Shaded reliefs have been used to visualize glacier dynamics previously (Wheate, 2012) and it is therefore obvious to apply the same to rock glaciers.

For the second animation, where the change in volume of the rock glacier is highlighted in color, difference images had to be calculated. To do so, a base layer had to be chosen. As already mentioned above, 1959 was chosen as the base layer. A difference image for each year compared to 1959 was calculated by subtracting the 1959 DEM layer from the DEM layer of the other year. That way a change in elevation and therefore volume was obtained (Rastner et al., 2016).

By the completion of these last steps, all the layers needed were ready to be further processed into 3D animations.

3.2.3 3D Animations and Color

To combine the shaded relief images as well as the difference images into an animation over time a few more steps were necessary. First, a mosaic dataset had to be created. To this mosaic, the rasters (shaded relief or difference images, respectively) of all the years were added. In the attribute table, two new fields were added: start and end time. It was necessary that both, a start and end time, were added, even though the raster layers only represent a single day of a specific year. The reason for that is that otherwise the animation would only show data on this single day and all the time in between the animation would be empty. Therefore, this timespan was created. The start and end times were determined by calculating the day in the middle of two years with data, assuming the aerial images used to compute the DEMs were taken on the 1st of January of every year. This is a simplification and does not represent the actual dates on which the aerial images were taken. For example, for the year 1978, the start time was 04.07.1975 as this was the day in the middle of 01.01.1973 and 01.01.1978 and the end time was 02.07.1980 as this was the day in the middle of 01.01.1978 and 01.01.1983.

Once the time variable was added to the mosaic, a time-slider appeared which allowed having a first look at how the rock glacier changed over time, still in a two-dimensional setting. At this point, the shaded relief mosaic was ready to be converted into 3D and into an animation. The difference images still needed one more step: color. The color of the difference images shows the change in height in meters. From this, information about the volume changes of the rock glacier can be derived.

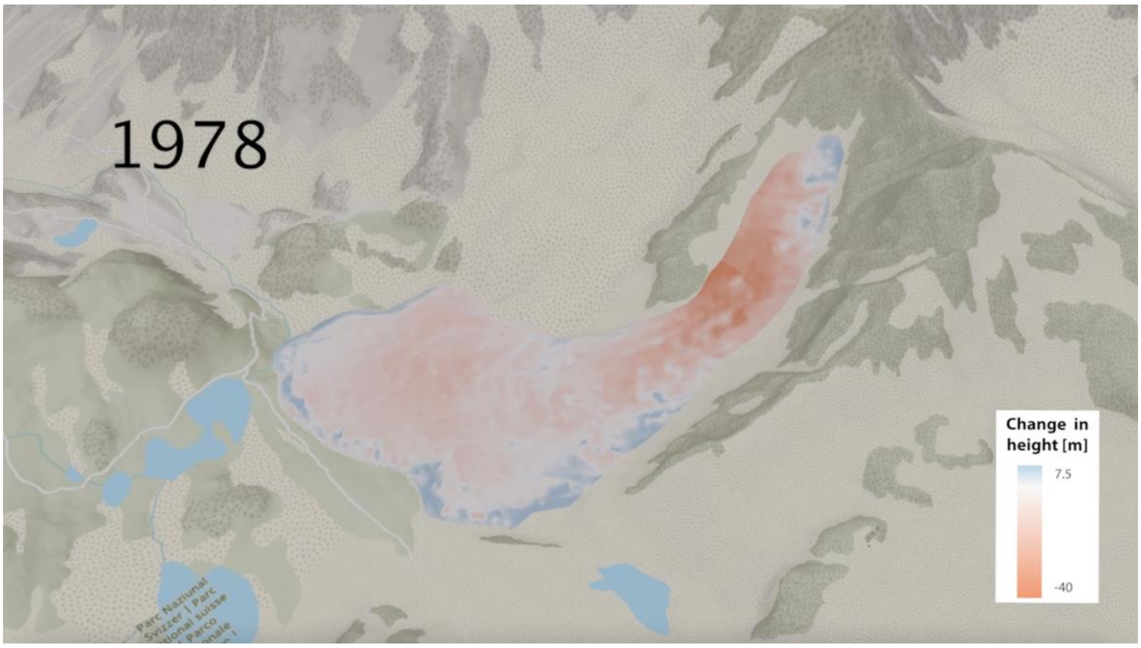
The color map visualizing the DEM differences is a diverging color scheme. As the red-white-blue diverging color map has shown itself to be effective (Moreland, 2009), this one was chosen. The exact colors were extracted from ColorBrewer. That way it was also possible to ensure that the color map was CVD-friendly. In ArcGIS Pro, the difference images could then be colored in a continuous color map having the values from ColorBrewer as minimum, midpoint, and maximum. There was one problem with that. By default, ArcGIS Pro maps the color map from the global minimum to the global maximum. The global minimum was -40 and the global maximum 7.5. Hence, in the color map, -40 and 7.5 were assumed to be equally far away from the midpoint, meaning that the values of the color map between -40 and 0 as well as 0 and 7.5 were stretched equally. From chapter 2.3.2, it is known that this leads to misperceptions of the data. Therefore, it was important to adjust this manually.

After the mosaic with the difference images was colored, both mosaics were ready to be converted into 3D. Using ArcGIS Pro this is quite an easy procedure as this is possible by simply clicking on the convert to 3D scene button. Creating an animation in ArcGIS Pro from the 3D scenes can be done with the animation tool. For an animation through time, only two keyframes had to be created. The first one shows the rock glacier in 1959 and the second one shows the rock glacier in 2003. Then the software automatically animates through time and the animations could be exported in .mp4 format. It was made sure that both animations have the same zoom level and angle so that there is no difference between them that could affect the outcomes of the experiment. The duration of the animations was also set to precisely the same time (30.1 seconds). The animation outputs were then further edited in Adobe Premiere Pro (version 22.5). In both animations, the display time of the start and end year had to be extended as in the animation from ArcGIS Pro they are only visible for a split second. The final length of the animations is 35.3 seconds. To both animations, a field showing the year which is currently displayed in the animation was added. For the animation with color, a legend explaining what the colors in the animation imply was also added. Figure 3.4a and 3.4b show the animation frame of the year 1978 as an example with color (a) and without color (b).

While it is common to include a directional indicator (north arrow) and scale in cartographic displays (Wilkening et al., 2019), it was decided not to include them in the animations. The main reason for that is that they are not important to answer the questions in the survey and the additional information in the animation may distract the participants from more important observations. Vicentiy et al. (2016) say that it is important not to overload visualizations with information that is not relevant and there is no need for the participants to know which way is north to answer questions about the animations. As there are no questions in the survey asking about absolute values (e.g., absolute change in volume [m^3]), a scale is also not relevant.

By the completion of these last steps, the animations were finished. Links leading to the final animations can be found in Appendix C.

(a)



(b)

Figure 3.4. Excerpt of the year 1978 from the animation with color (a) and without color (b).

3.3 Online Experiment

To assess the representation of rock glacier dynamics in three-dimensional animations with and without a diverging color scheme, an online experiment was conducted. This experiment was in the form of an online survey. In this section, the experiment design is explained in detail. The first part focuses on the experiment design itself. In the second part, the structure of the survey is described as well as the different types of questions that are a part of it. The third part of this section explains how the correct answers for the survey were calculated from the DEMs in ArcGIS Pro. In the final part of this section, the execution of the survey is described.

3.3.1 Experiment Design

The goal of the experiment is to obtain results that answer the research question and help to investigate the supporting hypotheses. In the setup with the two animations, there is one independent variable: color. An independent variable cannot be influenced by the participant's behavior (Martin, 2008). The independent variable can only be manipulated by the experimenter. In the case of this experiment, there are two levels of the independent variable. Level 1 shows the rock glacier over time with no addition of color and level 2 shows the same rock glacier over time with the addition of diverging color scheme indicating the change in height.

Besides color as the independent variable, there are also three dependent variables. The dependent variables are what is measured in an experiment (Martin, 2008). The formulated hypotheses focus on the better (H1) and faster (H2) representation of rock glacier dynamics when there is a diverging color scheme indicating the magnitude of change as well as a stronger confidence of the participants (H3). Hence, these are the dependent variables in this experiment. The dependent variable *accuracy* is assessed by comparing whether the given answers are correct if the question is a single-choice question. For the estimation questions, it is assessed how close the estimated value is to the correct value. There is also an open question in the survey (more on the question types in subsection 3.3.2). The open question is not mandatory for the participants to answer. Therefore, and because it is a qualitative question, it is not part of the main evaluation but rather an additional information source that may help to interpret the results and draw conclusions for future research.

The assessment of the dependent variable *time* is divided into two parts. For one, the time for each question group will be measured in seconds as well as the time needed to complete the whole survey. Additionally, the number of times the animation was watched is also measured. That way it can be distinguished whether it took participants longer to answer a question because they needed more time to think or because watching the animation multiple times led to longer response times.

The third dependent variable, *confidence*, is measured by asking the participants after every question group how confident they are with their answer as well as with a question at the end of the survey asking for the participants' overall confidence. The measurement is on a 5-point Likert scale. The Likert scale was introduced to science in 1932 and is used to capture subjective information such as the feelings of participants in studies (Joshi et al., 2015). Asking for the confidence of a participant's answer is not only important for the assessment of the dependent variable but also for the interpretation of the other results, especially for the single-choice questions that do not offer an "I don't know"-option to a question. The confidence question prevents participants from simply guessing an answer but allows them to state how confident they are with the answer they have given (Wood et al., 2021). However, the confidence question is also useful for the estimation questions to assess whether the estimation given was a wild guess or whether the participants actually believe their answer to be accurate. In the experiment conducted, this can help to investigate whether the rock glacier dynamics are clearly better or worse represented in a group or if both groups answer the questions with low confidence. There might also be a correlation with the participants' confidence and the time they needed to complete the survey.

For the experiment design, it was of great importance to determine whether the participants were going to be exposed to both animations or just one. Martin (2008) describes two approaches and elaborates on their advantages and disadvantages. Exposure to both animations is known as a *within-subject design* while dividing the participants into two groups that each only see one animation is called a *between-*

subject design. The main advantage of a within-subject design according to Martin (2008) is that it requires fewer participants than a between-subject design because all participants see both levels of the independent variable and therefore, the number of participants needed in a within-subject designed experiment with two levels is half the number needed for the same experiment in a between-subject design. However, the within-subject design also has a very strong disadvantage: prior exposure. Once participants were exposed to a level of the independent variable, it will influence their perception of the second level. Applied to the experiment conducted for this thesis, this means that participants may answer questions differently when they, for example, first watched the animation with color and then the one without color because they remember things from the first animation that would not be visible in the second one. This is a huge advantage of between-subject design experiments compared to within-subject design experiments. However, there is also another disadvantage of between-subject designs besides the larger number of participants needed. To conduct an experiment in a between-subject design the participants must be split into two or more groups (depending on the number of levels of the independent variable). There is the possibility that there are individual differences in the group (Martin, 2008). This means that it could be that Group 1 would always answer faster and more accurately than Group 2, no matter whether the animation they watched was with or without color, just because the individuals in Group 1 have a quicker and better grasp than the participants in Group 2.

Nevertheless, it was decided to use a between-subject design for the experiment of this thesis, as the prior exposure effect may have a big influence on this specific experiment. To counteract the problem of the unequal groups in between-subject design experiments, it was decided to use randomization to determine which individual is assigned to which group. According to Martin (2008), this is an effective approach to remove a potential bias between groups. How exactly the randomization was applied is explained in the next sub-section (3.3.2).

With the dependent and independent variables determined, the way to measure them elaborated, and the decision to have the experiment in a between-subject design, the next step was to decide how to conduct the experiment. Martin (2008) lists several options on how to conduct an experiment without a typical setup. One of them is an online survey, which is what was chosen for the experiment of this thesis. The survey was designed with LimeSurvey (limesurvey.org), a web-based tool with many options when it comes to the creation of surveys. LimeSurvey also allows measuring the time a participant needed to answer all the questions in a question group as well as the overall time which is useful in terms of the assessment of the time variable.

3.3.2 Structure of Survey

The survey starts with a welcome message, explaining its procedure and purpose. In total, there are ten question groups in the survey of which most are mandatory. There are two open questions that are non-mandatory and can be skipped by the participants. The first eight question groups focus on the rock glacier dynamics with an emphasis on the change of volume of the rock glacier. This is followed by a group of sociodemographic questions and a group of general questions regarding the survey. By the completion of all the question groups, the participants were forwarded to the last page indicating the end of the survey and a message thanking them for their participation.

While the welcome message mainly explained what the survey is about and how it works, it also said that by proceeding with the survey, the participants acknowledge that their personal data would be collected and may be published for scientific purposes. The participants were also informed that they would remain anonymous and that their names would not appear in any published work.

After the welcome message, a hidden question is embedded. As the word "hidden" already implies, this question is not visible to the participants. The question is in the form of an equation expression. The function randomly assigns the participants to a number, either 1 or 2. The participants that received number 1, are directed to the questions about the animation with color, and participants that received number 2, are directed to the questions about the animation without color. As mentioned in the previous subsection, this randomization counteracts a potential bias between groups.

The first eight question groups are all structured similarly. Each group starts with a display of the animation at the top of the page. The animation was embedded in a way that it adapts to the size of the browser window. Additionally, the full-screen mode is allowed. The participants have the option to watch the animation or parts of it as many times as they want. Some questions ask about the dynamics of the rock glacier on a specific part of the rock glacier (lower, middle, or upper part). If that is the case, the animation is followed by an image of the rock glacier that indicates where these parts are on the rock glacier (Figure 3.5). The image was colored in color-blind-friendly colors according to ColorBrewer. To enhance the legibility of the text elements over the colored background, a white halo was added. Halos are a common cartographic tool to highlight text over colored background (Hermann & Carpentier III, 2006).

Following the animation (and image), there are either one or two questions regarding the rock glacier dynamics. At the end of each question group, except for group 8 (which is a non-mandatory question), there is a question asking for the confidence of the participant's answer and another question asking how many times the animation was watched.

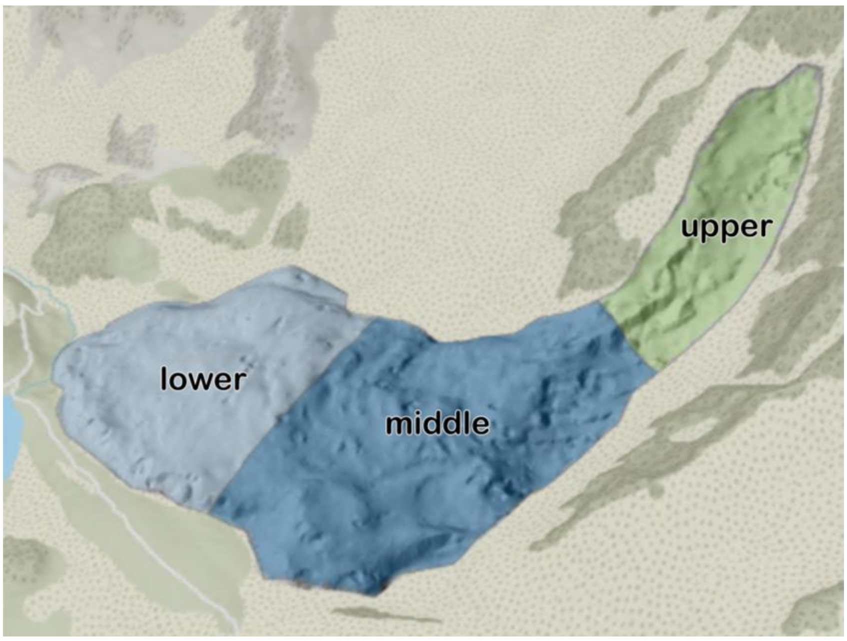


Figure 3.5. Image used in the survey (Group 2) to indicate where the lower, middle, and upper part of the rock glacier is.

The questions about the rock glacier dynamics are either single-choice, estimation, or open questions. The first five question groups consist of single-choice questions of four different types. The first question group focuses on the more general perception of the rock glacier dynamics simply asking the participants whether they see an increase or decrease in the rock glacier volume over the whole timespan. In the other five question groups with single-choice questions, there are questions asking the participants to distinguish in which part of the rock glacier (lower, middle, upper) the change was the highest or lowest. In other questions, the participant is asked to select a period. This means that the question asks the participants to compare two periods, for example, 1959-1969 to 1978-1988, and then decide in which period the increase or decrease was stronger. The fourth type of single-choice question is similar to the ones asking about the periods. The participants are asked to decide whether the rock glacier dynamics changed faster before or after a specific year. While this question technically also compares timespans, this question type focuses specifically on the rate of change of the rock glacier dynamics while the other questions focus on the increase and decrease of the rock glacier volume.

Question groups 6 and 7 are estimation questions. The participants are asked to estimate the change of the rock glacier volume as a percentage over the whole timespan as well as over a specific period. To indicate whether there is an overall increase or decrease, they have to add a plus or minus sign to their estimated number, respectively.

The eighth question group is an open question. The participants are asked to write what else they observe. This question is not mandatory and can also be skipped by the participants. This question group is not followed by a confidence question, unlike the first seven question groups.

As mentioned in the previous subsection, part of the time assessment is to ask the participants how many times they watched the animation per question group. Therefore, there is a question at the end of question groups 1-8 that only allows numerical input asking exactly that.

Question group 9 consists of sociodemographic questions. The participants are asked to answer questions regarding their age, gender, field of study or work, as well as their familiarity with geomorphological processes such as rock glacier dynamics. The age and field of study or work questions are open questions, the gender question is single-choice, and the question about the knowledge of geomorphological processes is designed with another 5-point Likert scale. Asking these sociodemographic questions is of great importance as it was shown in many different studies that they can have an impact on the results (Flandorfer, 2012). It was decided to ask these questions after the questions about the rock glacier dynamic in order to not distract the participants from the main topic of the survey.

The last question group technically consists of only one question. This question is again on a 5-point Likert scale asking the participants how difficult it was overall for them to answer the questions of the survey. This gives a general overview of the participant's ability to answer the questions with ease or whether it was hard for them and its part of the assessment of Hypothesis 3. This question is then followed by an open "question" leaving space for the participants to comment on anything they have still on their minds and also giving them the chance to give feedback.

The survey is then concluded with a thank you message. The complete survey with all the questions and answer options can be found in Appendix D.1.

3.3.3 Calculation of Answers

The correct answers for the questions in the survey were calculated in ArcGIS Pro with the Raster Calculator function. For the calculations, the DSMs reduced to the size of the rock glacier extent were used. For example, for the first question which asks whether there was an increase or decrease in volume over the total timespan over the whole area of the rock glacier, the DSM of the year 1959 was subtracted from the year 2003. Then the mean value of the raster dataset was calculated. The mean value was negative (-8.03m) and therefore there was a decrease in volume. A positive mean value would indicate that there was an increase in rock glacier volume. The same procedure was applied to question groups 2, 3, and 4.

When the questions asked to compare the different parts (lower, middle, upper) of the rock glacier, the DSMs were first clipped with shapefiles of the respective part and then subtracted from each other, and the mean was calculated from the output dataset. The shapefiles were originally self-drawn polygons such as the one that was used to clip the original DSMs to the rock glacier extent (subsection 3.2.2). As the polygons for the rock glacier parts were manually drawn, they are not very accurate in representing the exact border between the parts so that they are all equally sized. However, it was paid attention to make sure that the steepness correlates with the parts, meaning that the upper part is a lot steeper than the middle and lower parts and the lower part has the smallest inclination.

For question group 5, where the participants were asked to investigate the rate of change in m/a, the mean values of the two compared periods were calculated in the same way as for the previous questions. These were then simply divided by the number of years in the periods.

For the estimation questions which ask the participants to give a percentage, the mean values of the start and end years of the respective periods were calculated. The earlier year was then defined as 100% and the percentage value for the later year compared to the earlier year was calculated. The differences are very small (< 1%). Therefore, it was likely that the participants will over- or underestimate the change in percentage severely. However, as the correct answers for one of the estimation questions is a positive value (increase) and the other a negative value (decrease), it could still be assessed whether the participants have an accurate perception of the direction of change.

The correct answers for all questions as well as all the calculated values can be found in Appendix D.2.

3.3.4 Execution of Survey

Before the survey was sent to potential participants, a pilot run was conducted. Pilot experiments are important to detect problems within an experiment and solve them before the actual experiment (Martin, 2008). To do so the survey was made available to two people. It was made sure that one of them receives the survey with the animation with color and the other one the survey without color. The pilot participants then answered the questions in the survey as though they were participants in the actual experiment.

Both pilot participants did not encounter any problems in the survey and reported that the instructions on the individual questions were clear. Therefore, no adjustments had to be made. With the pilot run, it could also be tested how the responses are stored in LimeSurvey and in what formats they could be exported. It could also be ensured that the automatic time measurement from LimeSurvey per question group works.

As the pilot run did not reveal any issues, the next step was to distribute the survey. To do so, it was decided to send the link out via e-mail to the Department of Geography of the University of Zurich as well as to friends and family. It was important to not only include people with a geographic background as the goal was to create animations that are easily perceivable by anyone who is interested in the topic and not only the scientific community. To enlarge the range of people doing the survey it was also noted in the e-mail that everyone is welcome to share the link with other people they know. While online surveys allow to reach many people and gather information from them, the response rate may be rather low (Martin, 2008). By asking the possible participants to share the link chances rise that more people will answer the survey.

In the e-mail that was sent, it was also briefly explained what the survey is about, and prerequisites/restrictions were listed. As the survey should reach as many people as possible, there were not many. The only real prerequisite to participate is basic English knowledge as the questions as well as the instructions are in English. The participants were also advised to do the survey on a computer instead of a mobile device even though that would technically be possible as well. However, observing the rock glacier dynamics is easier on a larger screen.

After the survey was completed, the answers as well as the timestamps for each question group and the overall time were exported in a .csv-file for the analysis.

4 Results

The survey answers were evaluated with the statistics software R, version 4.1.2. In the beginning, a few outliers had to be removed from the data. For the estimation questions, a participant answered "-1'000'000%" change, which does not make sense as the whole rock glacier including its surrounding area would have had to have disappeared. It was concluded that this participant probably did not understand the question correctly. The rest of the participant's answers were included in the evaluation, only for the evaluation of the estimations questions they were excluded. There were also two outliers in the total time the participants need to complete the survey. Two participants had total times of 7'300 seconds (~ 2 hours) and 22'468 seconds (~ 6 hours). The third highest value for the total time was 2862 seconds (47 minutes). The two pilot runs needed approximately 30 minutes to complete the survey. It was therefore concluded that the participants that needed more than one hour to complete the survey were probably distracted or did not finish the survey at once. For the accuracy, the answers were included but not for the assessment of the time.

The data were evaluated with a variety of statistical tests. Depending on the data type and distribution (normal or non-parametric), a test was chosen. Hence, for the statistical comparison of means, either a t-test (normal distribution) or a Mann-Whitney U test (non-parametric distribution) was used. Further, multiple correlation models were created. As none of the data used for the correlation models was normally distributed, Spearman's ρ for non-parametric data was used to assess the relationship between two variables. For all comparison of means tests as well as the correlation models, a significance level of $\alpha = 0.05$ was defined. Hence, if the p-value of a statistical test was below 0.05, the result was statistically significant.

This chapter first describes the participants and gives information about their demographics. Following are a subsection focusing on the accuracy of the answers (H1) and a subsection focusing on the time of the answers (H2). The focus there lies on the statistical differences between the two groups. The fourth subsection focuses on the confidence of the participants' answers (H3). Lastly, the answers to the qualitative question of the survey (question group 8) as well as the optional further comments at the very end of the survey are described and assessed.

4.1 Sociodemographics of Participants

In total, 86 people participated in the online experiment and completed the survey (Table 4.1). Through the randomization process, 44 participants were assigned to Group 1 (animation with color) and 42 to Group 2 (animation without color). The age ranges from 19 to 62 in Group 1 and from 21 to 68 in Group 2. The mean age in Group 1 is 28.9 and in Group 2 29.7. Hence, the age distribution is similar in both groups. Regarding gender, the groups are also similar. There are 14 females and 29 males as well as 1 non-binary person in Group 1 and 17 females and 25 males in Group 2. In Group 1, there are 32 participants with a background in geography and 12 with no background in geography. In Group 2, it is slightly more balanced with 24 participants with and 18 participants without a background in geography.

Table 4.1. Summary of the sociodemographics of the participants of the online experiment.

	Group 1	**Group 2**	**Total**
Age (mean)	28.9	29.7	29.3
Gender			
Female	14	17	31
Male	29	25	54
Other	1	0	1
Background in geography			
Yes	32	24	56
No	12	18	30
Total	44	42	86

While there might be some small differences in the sociodemographics between the two groups, none of the dependent variables seemed to be influenced by one of these parameters. Therefore, these differences had no influence on the outcome of the experiment and the obtained results.

4.2 Accuracy (H1)

The first hypothesis focuses on the accuracy of the participants' answers. As it would be significantly more complicated to assess the single-choice and the estimation questions together, it was decided to evaluate them separately. For the single-choice questions, a correct answer received a score of 1 and a wrong answer a score of 0. By that, it was easy to evaluate the percentage of correct answers per group for every individual question. As question group 2 had two subquestions, each subquestion was only assigned 0.5 points instead of 1. To assess the overall accuracy of the single-choice questions a combined score was calculated. This means that for every correct answer a participant gave, they earned +1 for their score (or +0.5 for question group 2). Hence, the maximum score that participants were able to receive was 5, and the minimum score 0.

For the estimation questions, it was not possible to simply classify answers as correct or wrong as it was with the single-choice questions. Instead, the deviation of the estimated value to the actual value was calculated. The closer the estimated value was to the actual value, as the more accurate the answer was evaluated. Again, the questions were looked at individually first, and then they were combined into a single value by calculating the summed deviation a participant had for both questions. Additionally, it was evaluated whether the sign (+/-) of the estimated answer was correct or wrong. That way it was possible to assess whether participants correctly assumed the direction of volume change. This was done in the same way as with the score (assigning 1 to a correct sign and 0 to a wrong sign).

4.2.1 Single-Choice Questions

The results of the single-choice questions show that there are differences between the two groups regarding accuracy. Figure 4.1 shows that in 4 out of 5 single-choice questions more participants in Group 1 answered accurately than the participants in Group 2. Only question 3 was answered more accurately by Group 2.

In Group 1, every participant was able to answer the first question correctly. There is no question that was answered incorrectly by an entire group. Generally, question 5 seemed to be harder than the other questions as the percentage of correct answers is low in both groups. However, the differences between the groups are not statistically significant for every question (Table 4.2). Out of the 5 single-choice questions, 2 have statistically significant differences: question 1 and question 4. They are also the two questions with the biggest visual differences in Figure 4.1. In both cases, the p-value is even below 0.01 which suggests that the difference is very strongly significant. Both times it is Group 1 that performed better and answered the question more accurately.

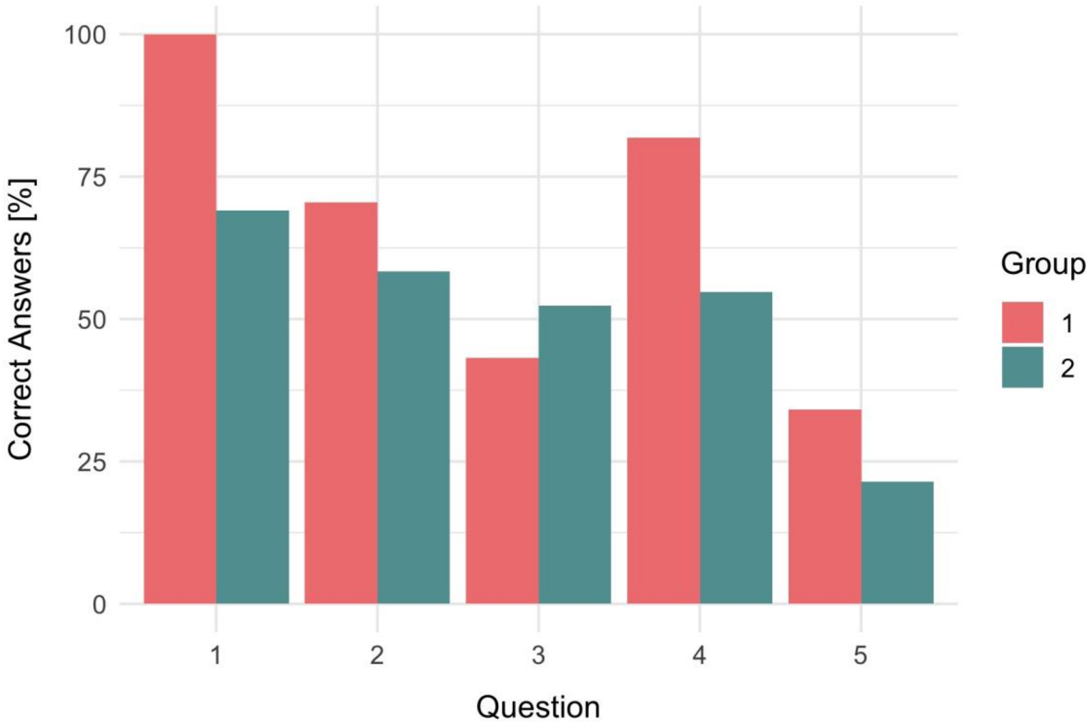


Figure 4.1. Percentage of correct answers per group per single-choice question.

Table 4.2. Percentages of correctly answered single-choice questions per group and the p-values of the Mann-Whitney U significance test.

Question	Correct answers [%]		p-value	Statistical significance
	Group 1	Group 2		
1	100	69.04	7.031e-05	Statistically significant
2	70.45	58.33	0.09807	Not statistically significant
3	43.18	52.38	0.3988	Not statistically significant
4	81.81	54.76	0.007333	Statistically significant
5	34.09	21.42	0.1952	Not statistically significant

As already explained above, the single-choice questions were not only looked at individually but also combined to see whether there is one group that answered the questions more accurately overall. The score ranges from 1.5-5 in Group 1 and 0-4.5 in Group 2. Figure 4.2 visualizes the distribution of the achieved scores by the participants per group. It is well distinguishable that the median line, as well as the 75th percentile, of Group 1 is above the one of Group 2, indicating that Group 1 answered the questions more accurately. This difference also becomes evident when comparing the means, which are 3.30 for Group 1 and 2.56 for Group 2. Furthermore, the difference between the two groups is statistically significant with a p-value of 0.002937.

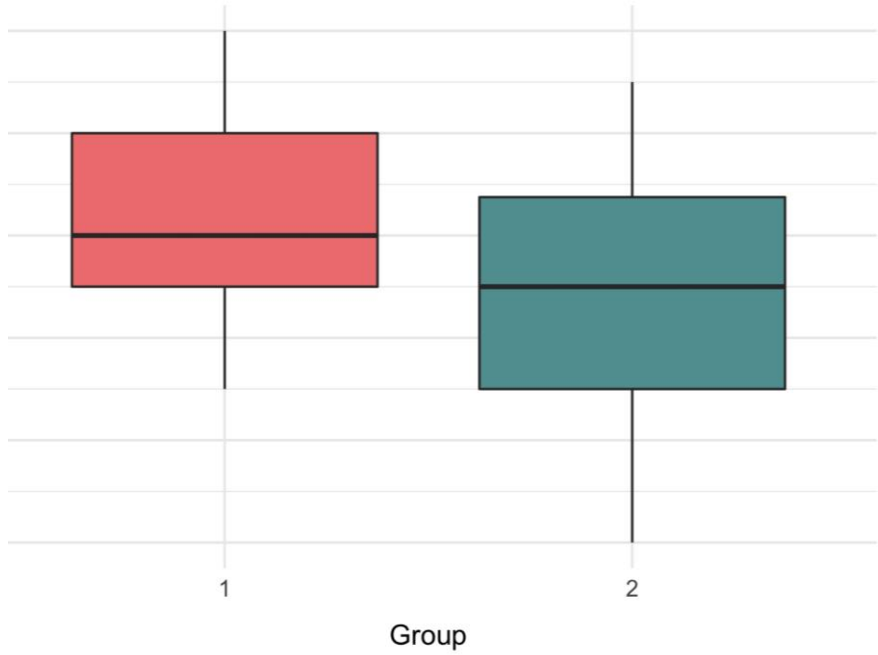


Figure 4.2. Score distribution of the single-choice questions per group.

4.2.2 Estimation Questions

The accuracy of the estimation questions was a bit trickier to assess as there was not a simple way to classify an answer as correct or wrong. When looking at the deviation of the estimated value from the actual value for the individual questions (Figure 4.3), it can be seen that both groups tended to overestimate the magnitude of change. Both groups generally underestimated the volume change in both questions. The dispersion is larger in Group 1 than in Group 2. Group 2 appears to have answered more accurately as the median line, as well as the 25th percentile line, is a lot closer to 0 deviation than the lines from Group 1. This characteristic becomes even more evident when looking at de boxplot of the combined deviation (Figure 4.4). The difference between the two groups seems to be large and Group 2 looks like it performed better in terms of accuracy than Group 1. The 75th percentile line of Group 2 is almost exactly at 0. The dispersion of the answers is different between the groups than it was when looking at the individual questions. However, Group 1 still appears to have a large dispersion.

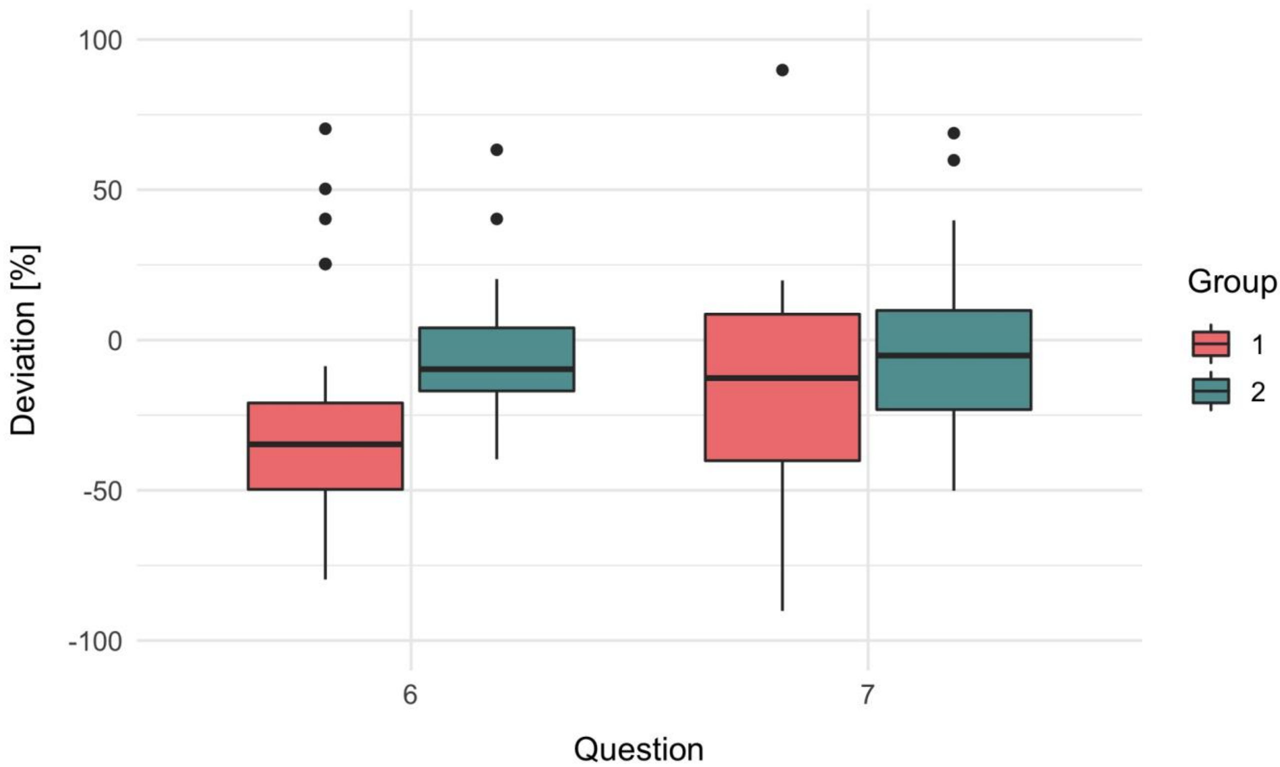


Figure 4.3. Deviation [%] of the estimated value to the actual value for questions 6 and 7. Note that the unit of the deviation is in % because the participants were asked to estimate the change in %. It is not the percentual deviation from the actual value.

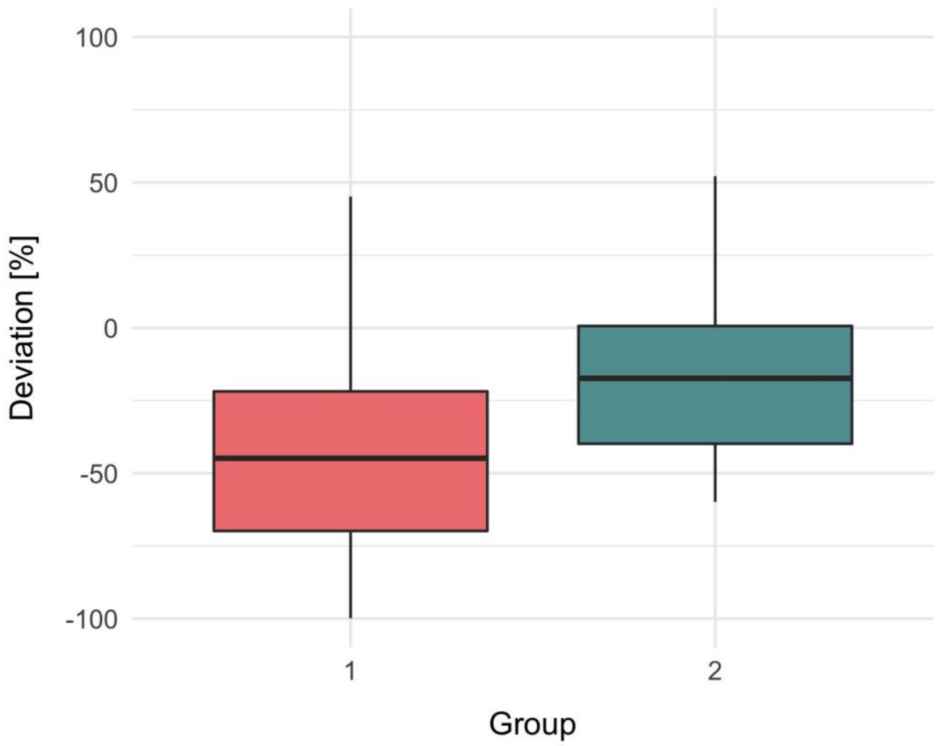


Figure 4.4. Combined deviation [%] of questions 6 and 7 from the estimated value to the actual value. Note that the unit of the deviation is in % because the participants were asked to estimate the change in %. It is not the percentual deviation from the actual value.

Unsurprisingly, the conducted significance tests confirm what was expected from looking at Figures 4.3 and 4.4. The p-values (Table 4.3) of the Mann-Whitney U tests show that the differences between the groups are statistically significant for the combined deviation as well as question 6.

Table 4.3. Mean deviation from the estimated values to the actual values per group and the p-values of the Mann-Whitney U significance test.

Question	Mean deviation [%]		p-value	Statistical significance
	Group 1	Group 2		
6	-29.30	-5.32	6.736e-07	Statistically significant
7	-19.61	-3.48	0.0695	Not statistically significant
Combined	-48.91	-8.80	2.21e-05	Statistically significant

Next to the deviation of the estimated value to the actual value, the correctness of the sign was also investigated. While the volume change is negative in question 6, it is positive in question 7. Hence, the sign was counted as correct if the participants' answers started with “-” in question 6 and “+” in question 7 respectively. While for both questions more participants in Group 1 than in Group 2 answered with the correct sign, the difference is not statistically significant (Table 4.4).

Table 4.4. Percentage of correct sign in the estimation questions per group and p-values of the Mann-Whitney U significance test.

Question	Correct sign [%]		p-value	Statistical significance
	Group 1	Group 2		
6	85.71	73.81	0.1792	Not statistically significant
7	61.90	59.52	0.8284	Not statistically significant
Combined	73.81	66.67	0.5728	Not statistically significant

In summary, while the single-choice questions were generally answered more accurately by Group 1, the estimation questions were answered more accurately by Group 2. While the participants in Group 1 answered with the correct sign more often than the participants in Group 2, the difference is not statistically significant.

4.3 Time (H2)

The second hypothesis of this thesis focuses on the time it took participants to answer the questions. To investigate this the time participants needed to complete the whole survey as well as for the individual questions were looked at. As already explained at the beginning of this chapter, two outliers had to be removed prior to the analysis. Their time measurements were also excluded for the assessment of the time of the individual questions and not just the overall time.

4.3.1 Overall

Looking at the time participants needed to answer all questions (Figure 4.5), it can be seen that the participants in Group 1 generally needed longer to complete the survey than the ones in Group 2. The dispersion in Group 1 is larger than the one in Group 2, especially towards the upper end. The median line of Group 1 as well as the upper and lower percentile line are above the ones from Group 2. The mean values are 918.47 s and 674.01 s for Group 1 and Group 2 respectively. A Mann-Whitney U test confirmed that this difference is statistically significant with a p-value of 0.01307.

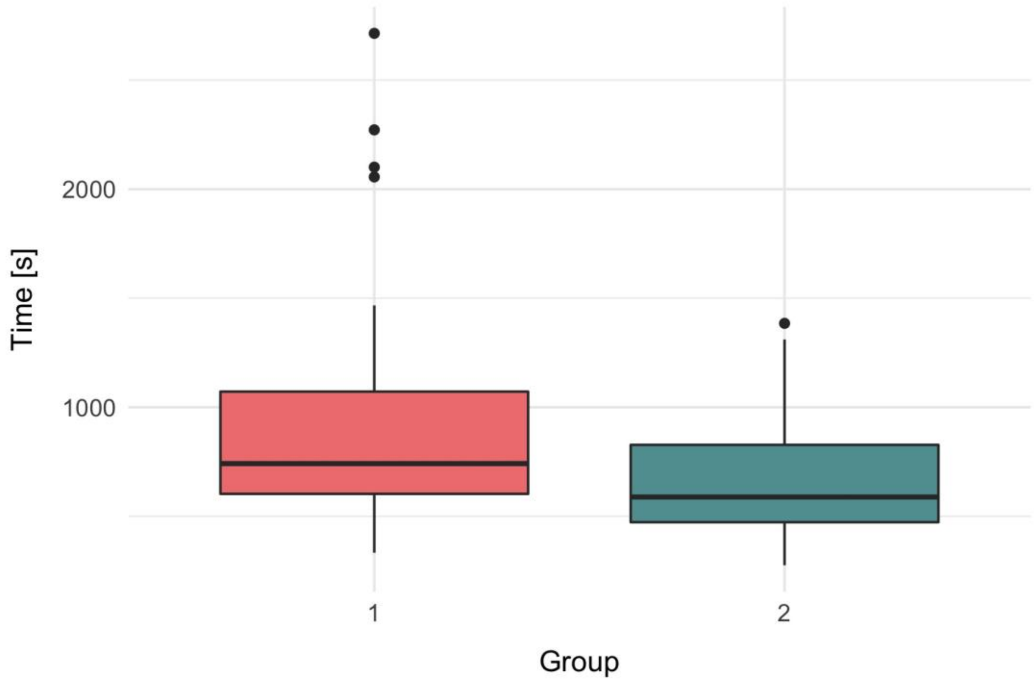


Figure 4.5. Overall time [s] per group participants needed to answer all questions.

4.3.2 Individual questions

Group 2 answering the questions faster than Group 1 is also visible when looking at the time participants needed to answer the individual questions (Figure 4.6). Group 2 answered every question faster except for question 1. However, the differences seem to be not that large for most questions. Questions 2, 6, and 7 appear to have slightly larger gaps between the two groups. The most eye-catching difference is the one in question 6, where Group 1 (159.61 s) needed on average more than double the time Group 2 (70.01 s) needed on average to answer the question.

As for the statistical significance of the differences in answering time per question between the groups a pattern becomes distinguishable. Group 1 needed statistically significantly longer to answer questions 6 and 7 with p-values of 0.000383 and 0.02611 respectively (Table 4.5). Questions 6 and 7 were the estimation questions. However, there is no statistically significant difference in answering time between the groups for the single-choice questions.

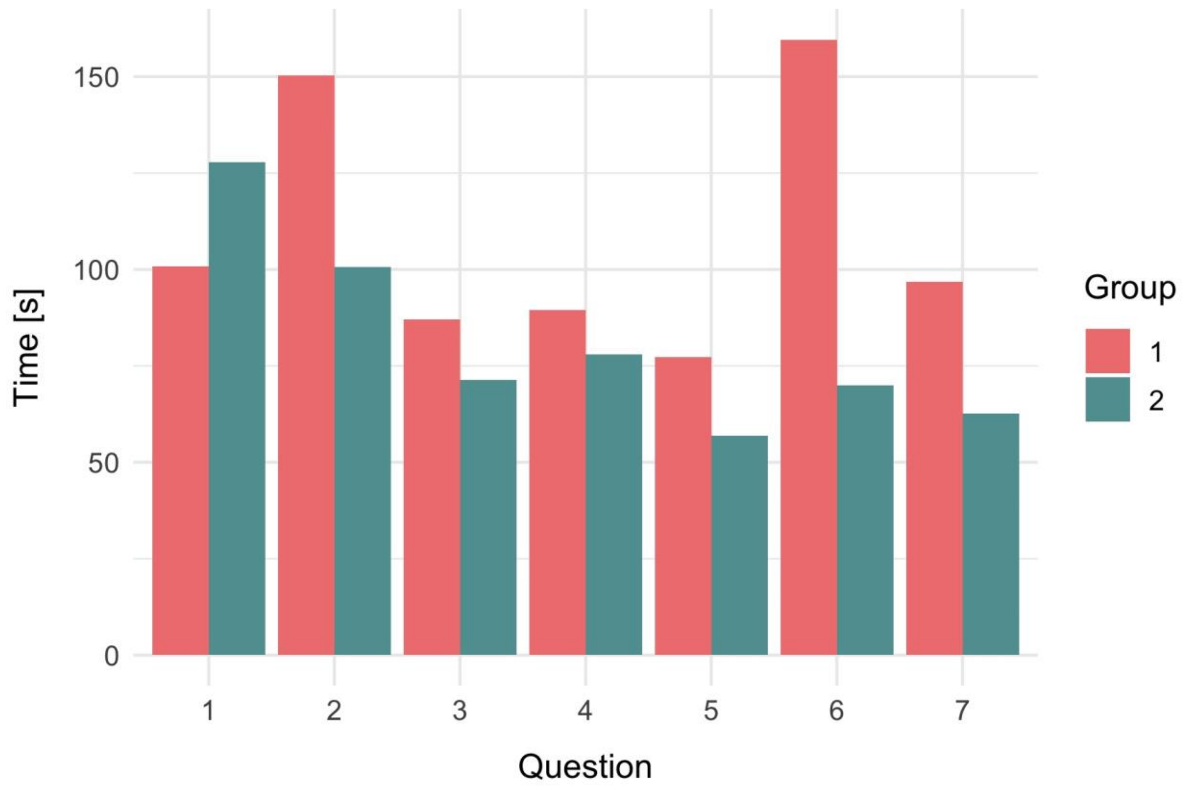


Figure 4.6. Time [s] participants needed per group for every individual question.

Table 4.5. Mean values of time needed to answer a question by group and p-values of the Mann-Whitney U significance test.

Question	Mean time [s]		p-value	Statistical significance
	Group 1	Group 2		
1	100.78	132.87	0.2198	Not statistically significant
2	150.36	101.27	0.3973	Not statistically significant
3	87.00	74.22	0.1024	Not statistically significant
4	89.45	78.02	0.1925	Not statistically significant
5	77.29	56.84	0.1786	Not statistically significant
6	159.61	70.01	0.000383	Statistically significant
7	97.73	62.72	0.02611	Statistically significant

4.3.3 Animation-Time Trade-Off

To assess whether it took participants longer to answer a question because they had to watch the animation more times to understand the process or because they needed more time to think, participants were asked at the end of each question group how many times they have watched the animation. The means for the times the animation was watched over the whole survey are 10.11 for Group 1 and 11.69 for Group 2. The difference between the groups is not statistically significant with a p-value of 0.2038.

To investigate the relationship between the number of times the animation was watched and the time it took for participants to complete the survey a correlation model was created with Spearman's ρ. The result (Figure 4.7) shows that there is a significant positive correlation between these two factors with a p-value of 6.8e-06 and an R-value of 0.47. This means that they indeed depend on one another. Hence, if a participant watched the animation more often than another participant it is likely that the first participant took more time to complete the survey than the second participant.

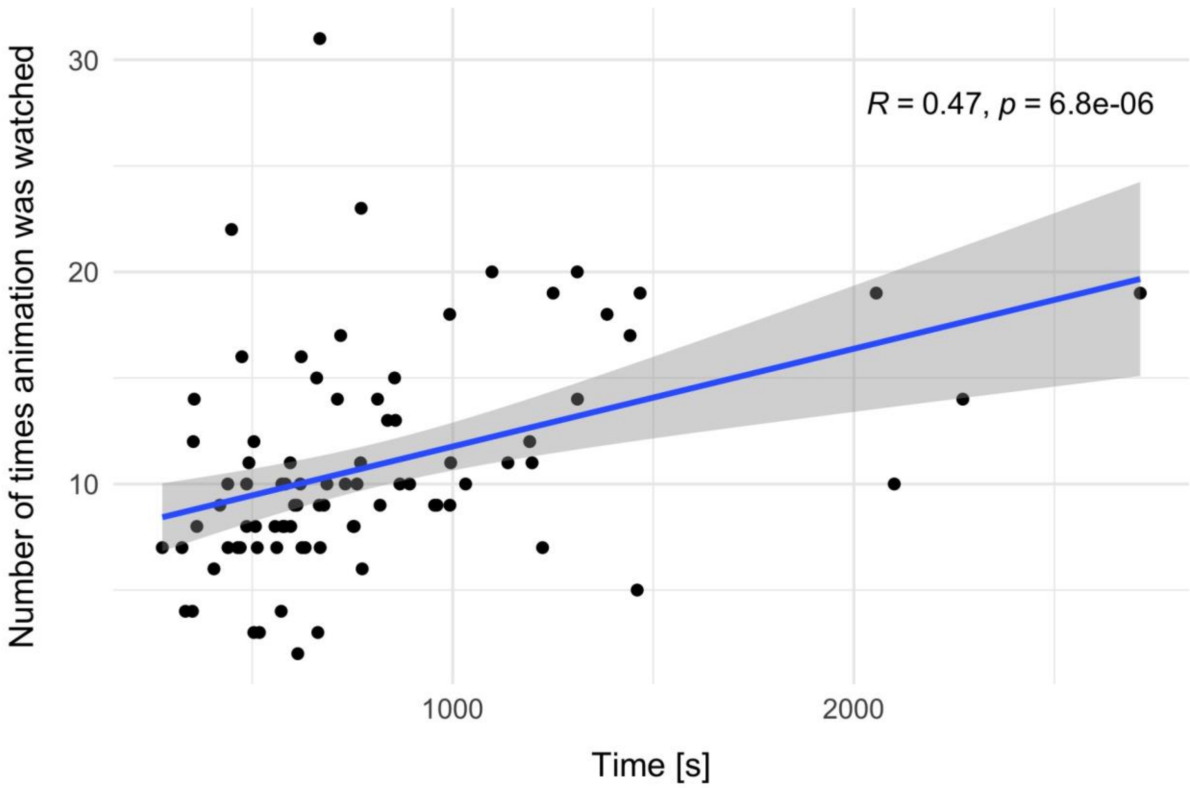


Figure 4.7. Correlation plot of the number of times the animation was watched, and the time needed to complete the survey, including a regression line.

It was also investigated whether there is a difference of the animation-time relationship between the two groups (Figure 4.8). For both groups, there is a statistically significant positive correlation with p-values of 0.00051 (Group 1) and 5.1e-05 (Group 2) and R-values of 0.5 (Group 1) and 0.6 (Group 2). The positive effect of the correlation is hence stronger in the individual groups than when looking at them in a combined setting.

Furthermore, it was compared whether the two correlation coefficients differ significantly by using the Fisher z-transformation. The z-statistic is -0.6344 and the p-value is 0.5259. It can therefore be concluded that there is no statistically significant difference between the two correlations even though the R-value is higher, and the p-value is lower for Group 2.

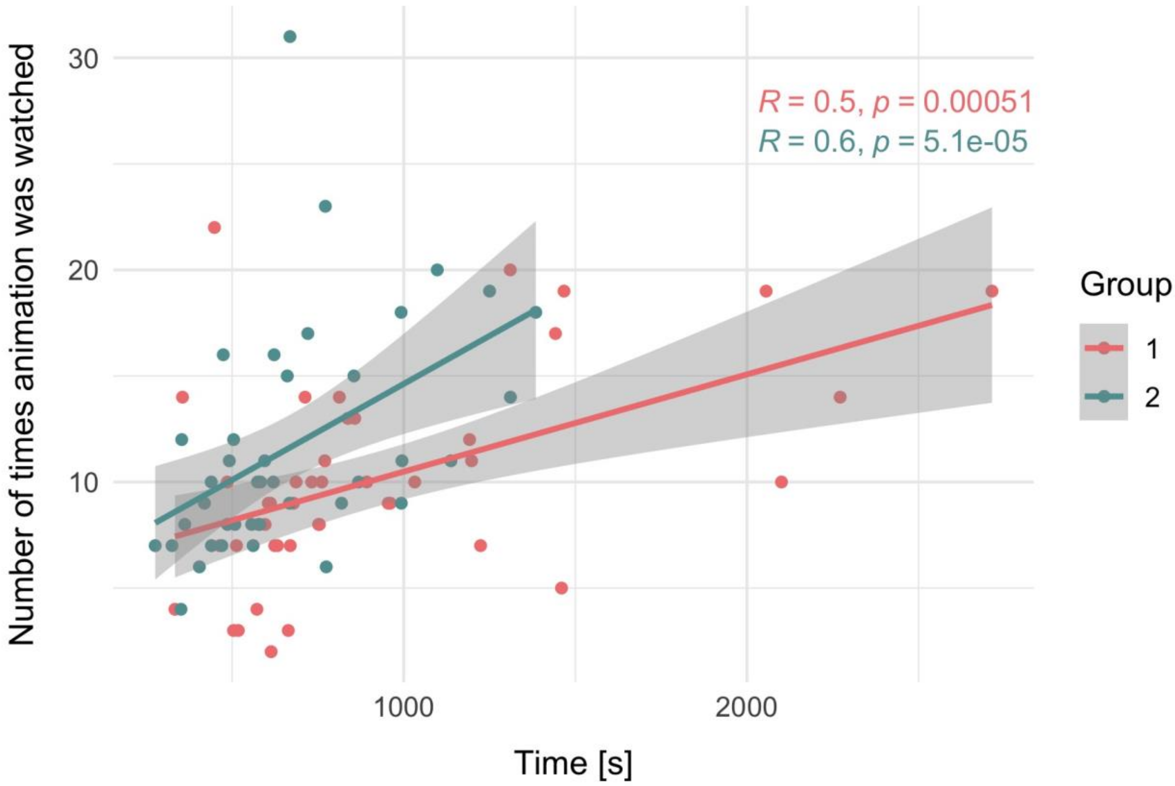


Figure 4.8. Correlation plot of the number of times the animation was watched and the time needed to complete the survey per group with the respective R- and p-values and regression lines.

4.3.4 Speed-Accuracy Trade-Off

As already mentioned in the introduction section of this thesis, the first two formulated hypotheses are contradictory to the speed-accuracy trade-off. Hence, having a look at the speed-accuracy relationship in the experiment may also give further insight into what the results of the survey might suggest for the hypotheses.

As the accuracy is difficult to assess for the single-choice and the estimation questions together, the trade-off with time was investigated separately for the two question types. However, for both question types, there is no statistically significant correlation between the accuracy of the participants' answers and the time needed to answer (Table 4.6). For both groups together, the accuracy of the estimation questions and the time even seem to have a negative relationship (R-value = -0.06), but, as already noted, it is not statistically significant.

Table 4.6. R- and p-values of the correlation analysis of the speed-accuracy trade-off per question type as well as per group and combined.

	R-value	p-value	Statistical significance
Single-choice			
Group 1	0.17	0.26	Not statistically significant
Group 2	0.043	0.79	Not statistically significant
Combined	0.14	0.20	Not statistically significant
Estimation			
Group 1	0.22	0.17	Not statistically significant
Group 2	0.021	0.90	Not statistically significant
Combined	-0.06	0.59	Not statistically significant

4.4 Confidence (H3)

Asking the participants about their confidence in an answer or in general led to a variety of results. For the overall confidence, it was possible to assess whether there was a group that felt more confident with their answers because the animation they watched left them with a stronger feeling of confidence. Furthermore, a correlation model with time as the second variable showed whether there is a relationship between the overall confidence and the time participants needed to complete the survey.

The confidence questions after each question group allowed to assess whether there is a correlation between the accuracy of the answers and the confidence and hence, whether the results include a lot of random accurate answers where participants simply guessed correctly or if wrong answers are often connected with low confidence.

4.4.1 Overall Confidence

Comparing the overall confidence between the groups reveals that Group 1 generally felt more confident with their answers than Group 2. This is well distinguishable in the boxplot (Figure 4.9) as the median line of Group 1 is at 2.5 while in Group 2 it is at 2. Furthermore, the variation in Group 2 is mainly towards lower confidence values. The 25^{th} percentile line is at 0, whereas for Group 1 the 25^{th} percentile line is at 2. For Group 2, the 75^{th} percentile line coincides with the median line, suggesting that it is also at 2.5. In Group 1 the 75^{th} percentile line is higher, at 3. Moreover, in Group 2 there are two outliers, one of which is at 4, whereas the upper whisker of the boxplot of Group 1 also reaches to 4. In both groups, there is an outlier at a confidence level of 5. The difference between the two groups is statistically significant with a p-value of 1.916e-05.

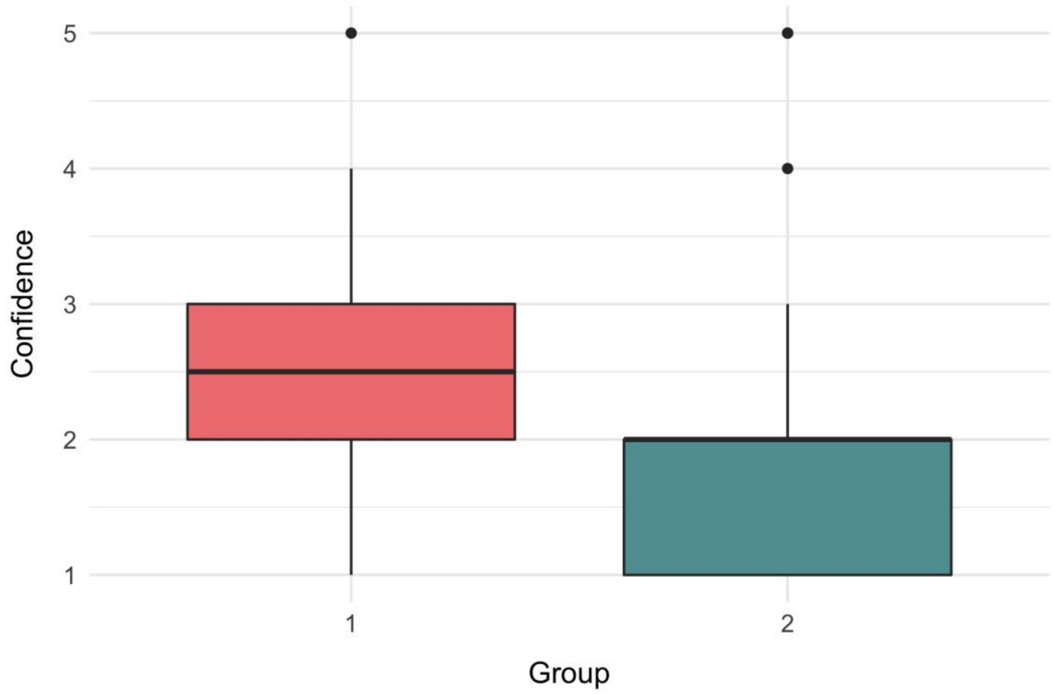


Figure 4.9. Boxplot of overall confidence levels (1 = low confidence, 5 = high confidence) per group.

Besides the difference in the overall confidence between the groups, the relationship between confidence and answering time was also assessed. Comparing the confidence with the time with no difference being made between the groups the R-value is 0.042

and the p-value 0.71. Hence, there is no statistically significant correlation between the two factors. Nevertheless, it was compared whether they may have a more significant relationship if the groups were considered. Figure 4.10 shows the correlation between the two, including a regression line. The confidence intervals around the regression lines are wide, suggesting large uncertainty. Just as the comparison of overall confidence and overall time, in general, had no statistically significant correlation, so the two variables had no relationship when divided into the two groups. Group 1 has an R-value of 0.052 and a p-value of 0.74. The values for Group 2 are -0.13 (R-value) and 0.41 (p-value).

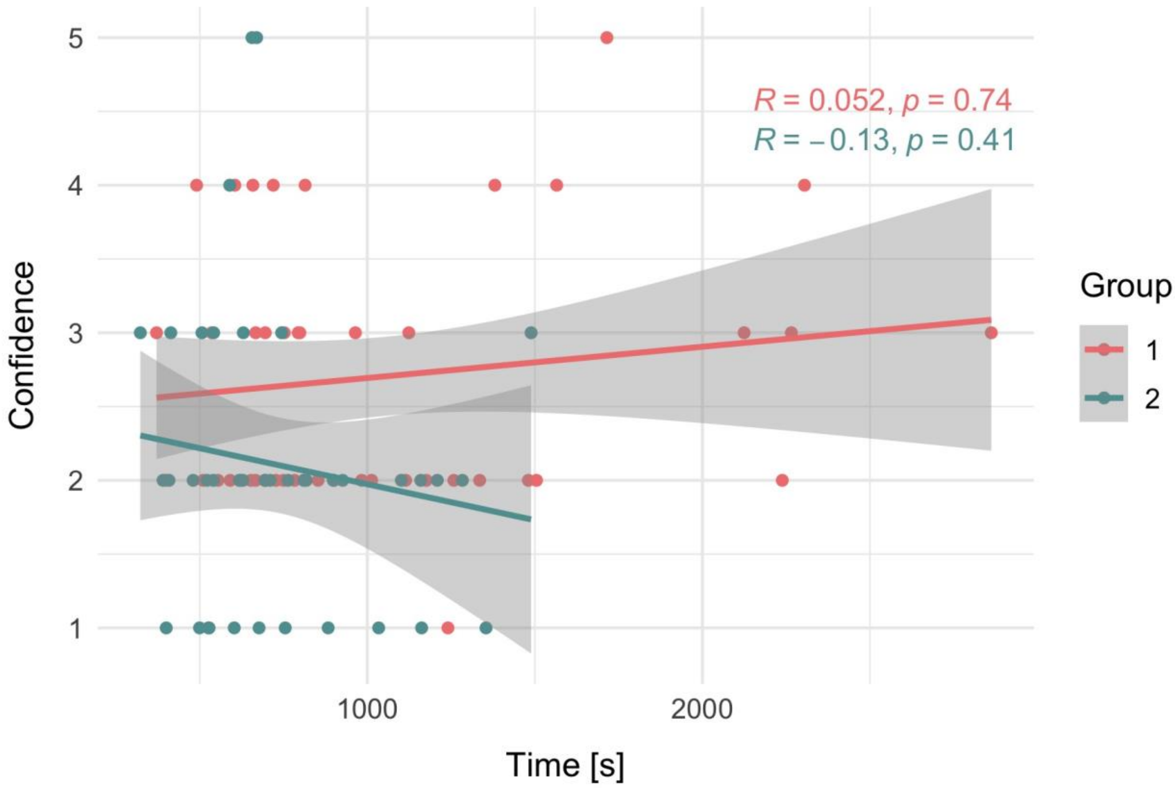


Figure 4.10. Correlation plot of the overall confidence of the participants and the time needed to complete the survey per group with the respective R- and p-values and regression lines.

4.4.2 Confidence-Accuracy Trade-Off

The confidence of the individual questions was combined with the accuracy (score and deviation). The confidence was summarized in the same way as with the score for the single-choice questions and the combined deviation in the estimation questions.

Hence, the summed confidence for the single-choice questions and the summed confidence for the estimation questions.

First, it was investigated whether there was a statistically significant difference between the confidence levels for the single-choice questions per group. While Group 1 had a mean confidence level of 18.61, Group 2 had a mean confidence level of 14.12. The difference between them is statistically significant with a p-value of 1.224e-07.

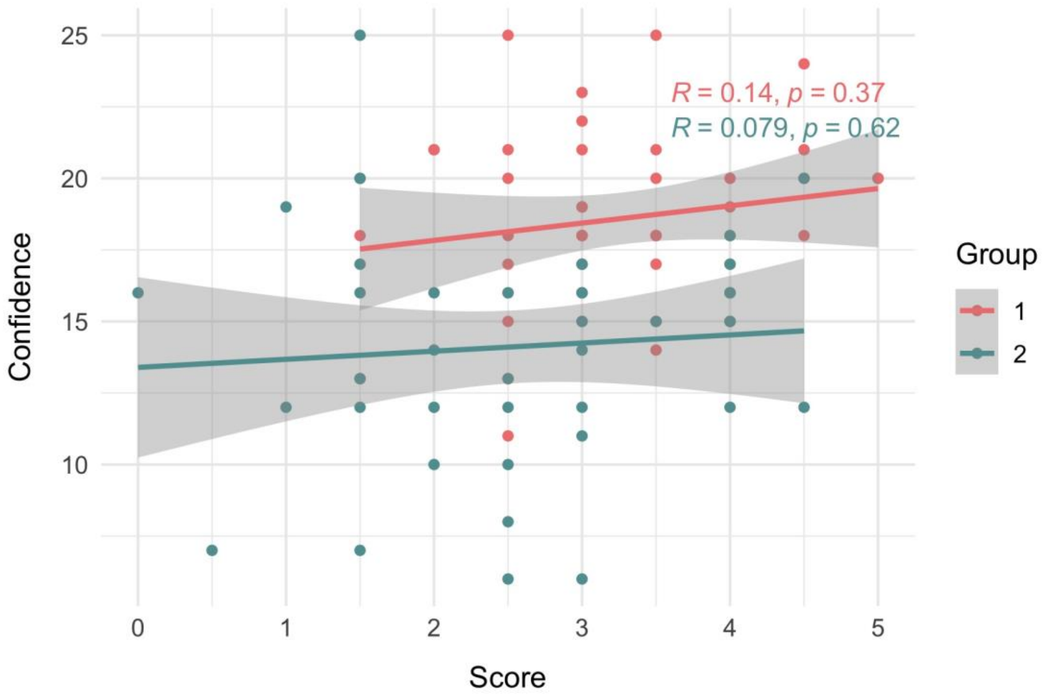


Figure 4.11. Correlation plot of the confidence and achieved score of the single-choice questions per group with the respective R- and p-values and regression lines.

While it was found that there is a difference in the confidence levels of the groups, there does not seem to be a statistically significant correlation between the confidence levels and the score achieved by the participants, considering the groups. Figure 4.11 shows the regression lines as well as the respective R- and p-values. The regression lines appear to be more or less flat and have a wide confidence interval. The R- values are 0.14 and 0.079 for Group 1 and Group 2 respectively. Both correlations are not statistically significant with p-values of 0.37 (Group 1) and 0.62 (Group 2). When looking at the distribution of the points in the plot this makes sense as they are widely distributed and most of them are not close to the regression line.

As well as for the single-choice questions, the first thing that was assessed was whether there is a statistically significant difference between the groups and their confidence regarding the estimation questions. The mean confidence level for the estimation questions for Group 1 is 4.86 and for Group 2 it is 3.93. The difference between the groups is statistically significant with a p-value of 0.0248. Nevertheless, the correlations between the confidence level and the deviation of the estimated value to the actual value are not statistically significant for both groups (Figure 4.12). The R-values are -0.077 for Group 1 and 0.087 for Group 2. The p-values are 0.63 and 0.58 for Group 1 and Group 2 respectively. As with the single-choice questions, the points in the plot are widely distributed for the estimation questions as well. The regression lines are only close to very few points and the confidence intervals are extremely wide.

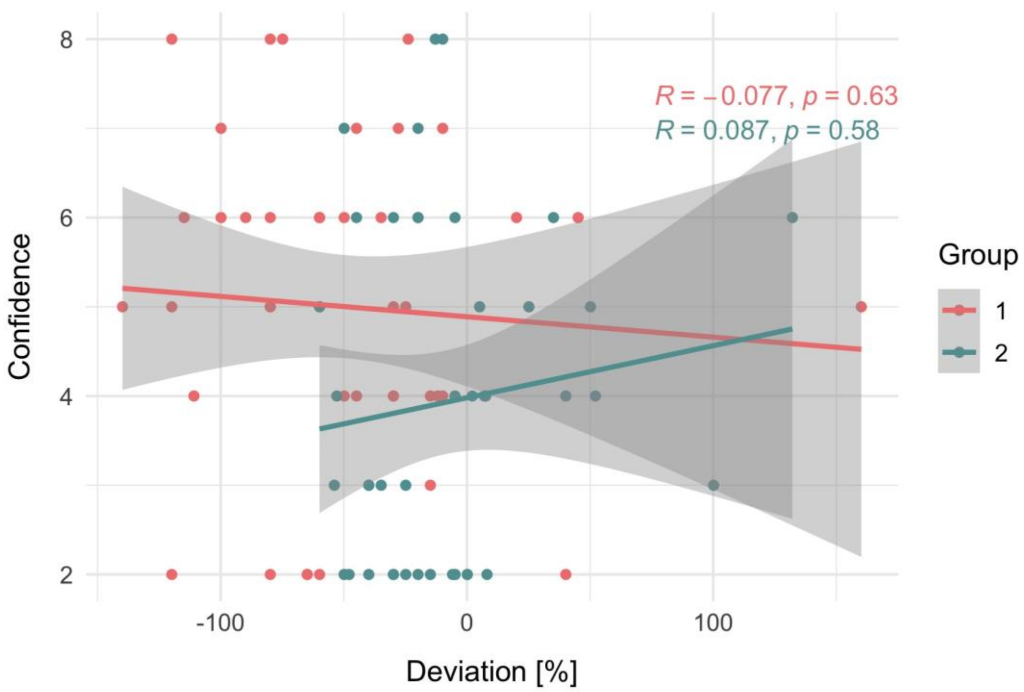


Figure 4.12. Correlation plot of the confidence and the deviation for the estimation questions per group with the respective R- and p-values and regression lines. Note that the unit of the deviation is in % because the participants were asked to estimate the change in %. It is not the percentual deviation from the actual value.

4.5 Qualitative Assessment

The qualitative assessment focuses on question group 8, in which the participants were asked whether they observed anything else in the animations and if so, what. Another point of focus is the further comments question at the very end of the survey. Both of these questions were optional and therefore, not every participant answered them. Additionally, some participants gave oral feedback after completing the survey as some of them were friends and family. This will also be mentioned in some cases, but it was not written down or further assessed.

Question group 8 was answered by 37 out of 44 participants in Group 1 and 40 out of 42 participants in Group 2. The further comments question was answered by 18 participants from Group 1 and 12 participants from Group 2. However, it has to be mentioned that not all comments were regarding the survey, but some were simply wishing good luck for the rest of the thesis. Subtracting those entries leaves 12 answers from Group 1 and 8 from Group 2.

First, the answers for question group 8 are described. The thing that a lot of participants of Group 1 noticed was that the upper part of the rock glacier records the highest volume loss. This was not mentioned by anyone from Group 2. Further, a variety of the participants from Group 1 noticed that there was no area change and no downhill movement visible even though that is what would be expected from what they know from their previous knowledge of glaciers and rock glaciers. In Group 2 only one person noticed this. In fact, 5 participants of Group 2 wrote that they noticed a downhill movement.

Further, over half of the participants of Group 2 noticed that there were changes in the resolution or the smoothness of the rock glacier and that there were some errors or holes in the shaded reliefs. It was also noticed by a participant of Group 2 that the time steps between the images are not always exactly 5 years.

In Group 1, 3 participants wrote that they made no further observations besides the volume change even after each of them having watched the animation another 4 times and many more that also did not see anything else but having watched the animation less often. There were 2 participants from Group 2 that also did not observe anything else after having watched the animation another 3 times.

While question group 8 focused explicitly on what else the participants observed besides the change in volume, the further comments question left room for many other things. A lot of the participants in Group 1 criticized that it was hard to assess the percentual volume of change without knowing the initial volume as a -40m change in height could be a very strong volume change or a very small one compared to the original elevation level. It was also mentioned connected to this that red is a strong color and the participant predicted that due overall, due to that, the participants would answer with very negative numbers. Further, it was criticized by some participants that the time steps between the represented years in the animation were uneven. What seemed to be unclear as well was whether the color of the change in height referred to how it was in 1959 or the previous frame (mentioned by 5 participants). One participant of Group 1 went as far as saying:

"This animation is a nice example of something looking fancy/pretty but providing limited information and ultimately being of very limited use."

Hence, this participant did not see any advantage in the addition of a color scheme. However, multiple participants from Group 2 suggested that "*color-coded elevations*" may improve the understanding of the volume of change of the rock glacier. It was also mentioned that having some sort of interactivity that allows to change the viewpoint of which the rock glacier is observed could have improved the understanding of the rock glacier dynamics.

Further, and this is what was also said in oral feedback, participants in Group 2 criticized that they hardly recognized anything in the animation. The participant who answered the further comments questions regarding this topic said they "*practically didn't see change in volume*". In the oral feedback, it was often stated that the participants were not able to distinguish any volume change and had to guess a lot of the answers.

5 Discussion

To answer the research question on how a diverging color scheme influences the perception of rock glacier dynamics, the experiment's results are analyzed and discussed in this chapter. Besides the research question, the hypotheses are taken up again and related to the findings of the study. Furthermore, it is discussed how the findings of this thesis relate to similar work that has been done in this field. In the last section, the limitations of the creation of the 3D animations as well as the experimental setup are described and analyzed.

5.1 Accuracy (H1)

The first hypothesis stated that Group 1 (color) would obtain more accurate results than Group 2 (no color). The results suggest that this is partly the case. For the single-choice questions, Group 1 indeed achieved better results. Looking at the individual questions, for 2 out of 5 questions, Group 1 performed significantly better and in 4 out of 5 questions there was no significant difference, but Group 1 still performed better than Group 2. Further, Group 1 achieved a significantly better score in the single-choice questions than Group 2. These findings suggest that, as expected, Group 1 answered the questions more accurately than Group 2. However, for the estimation questions, it looks very different. Group 1 performed significantly worse than Group 2 when looking at the estimation questions combined. Furthermore, Group 1 also performed worse in both estimation questions individually and in one of them significantly.

The results suggest that the animation with color gives a better general understanding of the rock glacier dynamics, but the extent of the dynamics is not as well presented as in the animation without color. A possible reason for that is that the general rock glacier dynamics were better distinguishable with color than by simply having the three-dimensional DEM as an information source. However, when the participants had to estimate the relative volume change of the rock glacier, they underestimated the value by far. In other words, they overestimated the magnitude of the difference. Hence, because the color of the rock glacier was tending more toward red than blue, the participants must have thought that the magnitude of change is a lot larger than what it actually was.

Some comments suggested that indicating the original volume or height of the glacier would have helped to estimate the change. This information was intentionally not given as the group with the shaded relief did not have that information either because it would have been no use for them. However, in retrospect, these comments make sense, and it probably would have led to a different result if the original height of the rock glacier was known. The information could have been given to both groups. It is well possible that the participants of Group 1 would have estimated more accurate values than the participants of Group 2 if the original rock glacier height was known. A 40m difference could indeed be very large or very small, depending on the original height. It would have to be further investigated whether giving this information would lead to more accurate answers of the group with color.

Additionally, in the comments section, it was also mentioned that red was a very strong color that would force more attention towards it than towards blue. This could be another reason why the magnitude of change was highly overestimated by Group 1 in the estimation questions. Another reason why Group 2 might have been better in the estimation questions than Group 1 is that, as also mentioned in the comments, they barely saw a change at all, which is why the estimated volume change may be closer to 0 than the estimation answers given by Group 1.

Nevertheless, even if Group 1 performed worse than Group 2 in the estimation questions, it was expected that Group 1 still performed better in determining the direction of change correctly. Yet, there was no significant difference between the two groups. However, the accuracy of the estimated direction of change was generally high in both groups, leading to the conclusion that both groups performed well in determining the direction of change and that the general direction of change can be determined with or without color as an additional information source.

All in all, the hypothesis about the enhanced accuracy with a diverging color scheme needs further research as no definitive answer can be proposed. The group with the diverging color scheme gave more accurate answers to the single-choice questions, indicating that there is an enhancement in the understanding of the rock glacier dynamics with the addition of a diverging color scheme. However, for the estimation questions, this is not the case and the diverging color scheme even led to less accurate results, which may be due to several reasons.

The findings, hence, partly agree with what has been found in previous studies. Chang et al. (2009) concluded in their study that the addition of a diverging color scheme on a three-dimensional model eases the understanding of what is being displayed. This seems to correspond with the results from the single-choice questions but not with the estimation questions. Therefore, there is no definite answer to whether the first hypothesis seems to be true or not and further research is required.

5.2 Time (H2)

The second hypothesis focused on the time the participants needed to complete the survey as well as to answer the individual questions. It was hypothesized that Group 1 would answer the questions faster than Group 2. However, the results show otherwise. Participants of Group 1 needed significantly longer to complete the survey than the participants of Group 2. This goes for both, the overall time as well as for each question group individually, even though not all differences in the individual questions were statistically significant.

A possible reason for this outcome could be that the participants of Group 2 indeed found it easier to perceive the rock glacier dynamics and thereby needed less time to complete the survey. This is also supported by the fact that, even though there is no statistically significant difference between the two groups and the number of times they have watched the animation, Group 2 watched the animation less often than Group 1, on average. The analysis of the animation-time trade-off showed that there is a significant positive correlation between the number of times the animation was watched, and the time needed to complete the survey. Hence, even if the difference between the two groups is not statistically significant, this might be an explanation for why it took the participants of Group 1 longer to answer the questions.

However, this brings up the question of why the participants of Group 1 had to watch the animation more times than the participants of Group 2. A possible explanation for this can be found by looking at the qualitative assessment. For one, the colored animation may have let the participants think there was more to see in the animation than what was actually visible, leading to longer answering times. This is shown in question group 8, where a lot of participants from Group 1 did not see any further differences but nevertheless watched the animation again multiple times. Some participants from Group 2 answered similarly but there were more in Group 1 with that answer. Further, as mentioned by one participant, the color may have distracted the participants by its pleasing aesthetics which may have led to watching it again with the hope of discovering something more when there was nothing more to see. Supporting this argument is that it was often mentioned by participants of Group 2 that they did not see anything in the animation. Therefore, it could be possible that the participants did not watch the animation more times because they thought that they were not going to see anything else anyway. However, this is just one possibility. It could also very well be that Group 2 answered faster because they understood the rock glacier dynamics faster. Further research would have to be put into that field.

The results are contradictory to the existing literature about the perception of time and the influence of color. For example. Hoarau & Christophe (2017) found out that the addition of color to ortho-imagery decreases the perception time of participants. Further, Chang et al. (2009) stated that less mental effort is needed with the addition of color, suggesting that the participants would need to put less time into understanding the information and therefore should answer faster.

When the hypothesis was first introduced in the introduction section it was mentioned that, combined with the first hypothesis, it is contradictory to the speed-accuracy trade-off which states that the more accurate an answer is, the slower the answering speed. The results would suggest that the outcome corresponds with the speed-accuracy trade-off for the single-choice questions as Group 1 answered more accurately than Group 2, but slower. However, the results of the investigation of the speed-accuracy trade-off show that there is no statistically significant correlation between the two factors. Hence, this is not the explanation for the unexpected outcome of the second hypothesis.

To summarize, Group 2 completed the survey faster than Group 1. This is not what was expected from the hypothesis. It may be possible that Group 2 indeed understood the rock glacier dynamics faster than Group 1 and hence, had shorter answering times. Another possibility is that the color led participants to think there was more to see than there actually was to see and that the animation without color did not show enough information to make the animation worth watching many times. Further research regarding this issue is needed. However, what is certain is that the speed-accuracy trade-off is not the reason for this unexpected outcome.

5.3 Confidence (H3)

The third hypothesis stated that the addition of color to the animation increases the participants' confidence in their answers. Looking at the overall confidence the participants of Group 1 were significantly more confident with their answers than the participants of Group 2. Hence, the outcome for the overall confidence is as expected. It also coincides with what was previously known, namely that additional information, such as color in this study, increases the participants' confidence in answering questions (Koriat et al., 1980).

What is also interesting is that not only was Group 1 more confident than Group 2 with their answers, but the confidence of Group 2 was extremely low with the median of the confidence level being 2 out of 5 (5 being the highest) and 75% of the participants of Group 2 answering 2 or less. This indicates that the participants were very unconfident and did not feel like the presented information was sufficient to answer the questions. This is supported by the fact that several participants of Group 2 suggested a sort of color coding for the elevation levels would have helped better understanding the rock glacier dynamics. This is exactly the sort of additional information that increases a participant's confidence that Koriat et al. (1980) mentioned.

While it seems to be evident that the addition of a diverging color scheme increases the participants' confidence in their answers, it could also be that other factors influence the participants' confidence. However, there is no significant correlation between the time the participants needed to complete the survey and their overall confidence. This means that spending more time thinking about an answer did not increase the participants' confidence. As there are no other measurable factors that may have increased the confidence, it was probably only the addition of color that influenced the participants' overall confidence.

Not only the time did not correlate with the confidence but there was also no significant correlation between the confidence of the answers to the individual questions and their accuracy. This means that even though the confidence was higher in Group 1, it did not have anything to do with whether an answer was correct or not. For the estimation questions, Group 1 even had a negative correlation between confidence and accuracy. However, it was not statistically significant. A negative correlation would imply that the more confident a participant is with their answer, the less accurate it is. This is certainly not what was wanted to be achieved with the addition of color to the animation. In an ideal case, there would be a positive correlation between the two variables, suggesting that the more confident a participant is, the more accurate the answer. A possible explanation for this outcome is similar to what has already been discussed above, namely that red is a very strong color, letting the participants think that there must be a very strong decrease in volume and at the same time increasing the confidence of the participants because red is so dominant in what they see. Again, as suggested by several participants, knowing the initial glacier volume may have helped to better perceive the proportions of the rock glacier and hence, the rock glacier dynamics.

As expected, the confidence of the participants in Group 1 was significantly higher than the one of the participants in Group 2. The difference in confidence between the two groups can only be related to the addition of color and not to other factors, such as answering time. However, the confidence does not correlate with the accuracy of the answers, indicating that these two variables also do not have a relationship.

5.4 Limitations

There are several limitations to the chosen approach. Some of these limitations already occurred during the creation of the 3D models that were later combined into the two animations. For one, the same outline polygon of the rock glacier was used for all 3D models for simplicity reasons. However, in reality, rock glaciers advance or retreat and maybe sometimes change their shape (Frauenfelder & Kääb, 2000). Hence, to obtain a more accurate model of the rock glacier, a new polygon would have to be drawn for each year. Additionally, there were shadows in some of the aerial images depending on the time of the day they were taken. These shadows sometimes made it hard to distinguish the border of the rock glacier and had to be manually corrected in the point cloud, which may have led to small deviations from the actual rock glacier.

A similar limitation is that the interpolation between the individual 3D models was conducted linearly even though, generally, rock glaciers do not retreat nor advance linearly (Rastner et all., 2016). Additionally, it was decided to use all the years with the original data in the animation which led to uneven time gaps between the years represented in the animation which was disturbing for a few participants. Ideally, there would be aerial images that allow the calculation of 3D models with photogrammetry for each year in the animation. This would solve the problem of untruthful interpolation and the uneven time steps.

Further, the calculated DEMs were based on a lot of manual work which has led to a variety of unrealistic values in the DEMs. For example, the maximum elevation in the DEM for 1959 was far too large and had to be corrected but probably still did not represent the actual value. Further, the values were sometimes unrealistic in general. The maximum decrease of the rock glacier thickness was around 40m which is very improbable regarding what is known about rock glacier dynamics (Steinemann et al., 2020).

All limitations mentioned until here were not relevant to the experiment itself as the experiment focused on the *perception* of the dynamics and hence, the represented dynamics did not have to be truthful. While the goal of this thesis focused on investigating whether the addition of a diverging color scheme enhances the perception of rock glacier dynamics in a visualization, this is only a part of finding an ideal visualization method for representing these dynamics in a way that they are easily understandable by everyone with an interest in the topic. Hence, there is a need for finding a solution for these issues in order to truthfully visualize the dynamics of a rock glacier.

Next to the limiting factors in the creation of the animations, there were also a few limitations in the experimental setup. The most important one is that the experiment was conducted in an uncontrolled setting. This uncontrolled setting brought a variety of connected limitations with it. For one, there is an issue with the time measurement. There are a few participants that needed a very long time to complete the survey which led to the conclusion that they did not finish the survey at once and left it running while doing other things in between. Additionally, in some cases, all but one question was answered within a similar time range and then one answer took the participant more than half of the total time they needed. They were probably distracted. In a controlled setting, this would not happen as the participants would only focus on the experiment. Further, the participants had to specify themselves how many times they have watched the animation after each question group. It is possible that some participants did not count correctly or if they only watched parts of an animation again, they did not know if that counted as another time or not. In a controlled setting, the executor of the experiment could make sure that the number is counted correctly and that partly watched animations are always counted in the same way.

Moreover, the known limitations of between-subject designs applied in this experiment as well. Martin (2008) explained that the main disadvantages of between-subject designed experiments are that a larger number of participants is needed and that there are individual differences between the groups. The number of participants was not really a limitation in this experiment, as enough people willing to participate were found. However, the second point may have been an issue. From the summary of the participants' sociodemographics, it is known that the groups (at least sociodemographically) are similar to each other. Nevertheless, it was not investigated whether the sociodemographic factors may have an influence after all. Therefore, further research regarding this topic is required.

Further, most participants were from the Geographic Department of the University of Zurich. Most of them have some knowledge about rock glaciers as well as the basic visualization principles. This was technically not limiting but may still have influenced the outcome of the experiment.

Several aspects were not really limitations to the experiment but could still be improved and thereby enhance the visualization of rock glacier dynamics. The one thing that was criticized the most by the participants of Group 1 is that the original volume or height of the rock glacier was not known. With this information, the result may have differed significantly. Another point mentioned is that having a look at the rock glacier from different viewpoints with some sort of interactivity may have helped to better understand the rock glacier dynamics. Interactivity is also one of the three principles a visualization should follow in order to make geomorphological processes and phenomena more easily comprehensible listed by Tateoisan et al. (2014). The influence of these two factors would have to be further investigated as it seems promising that they might have a positive influence on the perception of rock glacier dynamics.

In summary, there were some limitations to the approach. However, most of them did not affect the outcome of the experiment strongly but could merely enhance the visualization of rock glacier dynamics in general. The experiment was successful despite these limitations and could be executed without further issues.

6 Conclusion and Future Work

The goal of this thesis was to find a way to enhance the perception of rock glacier dynamics in 3D animations in a way that these dynamics become more easily understandable to everyone with an interest in them and not only to people with scientific knowledge about rock glaciers. As the addition of a diverging color scheme seemed to be a promising step in that direction, the research question of how a diverging color scheme influences the perception of rock glacier dynamics in 3D animations was formulated. To investigate this research question, three hypotheses focusing on the answer accuracy, answer time, and answer confidence were proposed. All three hypotheses suggest that the group with the animation with the diverging color scheme (Group 1) performs better than the group without the addition of color (Group 2).

In terms of answer accuracy, Group 1 indeed performed better for the single-choice questions than Group 2. However, for the estimation questions, Group 2 answered more accurately. Hence, there is no definite way to either reject or fail to reject the hypothesis. For the answer time, Group 1 performed significantly worse than Group 2 for the overall time as well as for the time of the individual questions. This outcome was not what was expected, which could be for several reasons. Hence, the second hypothesis had to be rejected. The third hypothesis, focusing on the answer confidence, has failed to be rejected as Group 1 had significantly higher confidence than Group 2. Hence, the third hypothesis is probably true.

Because not all hypotheses could definitely be rejected or failed to be rejected, several issues have to further be investigated in order to find an ideal visualization for rock glacier dynamics in 3D animations. For one, it should be further investigated why the diverging color scheme helped to answer the single-choice questions but achieved the opposite result for the estimation questions. A possible reason could be that there was no information about the initial rock glacier height given to the participants, which many said may have helped with the interpretation. Therefore, another experiment where this information is given should be conducted in order to find out whether this has an impact on the outcome. Further, it should also be investigated why the participants of Group 1 were slower in completing the survey than the participants of Group 2 as the evidence from prior research suggests otherwise. The results of the confidence of the participants coincide with previous research. However, what still could be further investigated is why there is no significant correlation between the answer accuracy and confidence.

Besides optimizing the application of a diverging color scheme on 3D animations of rock glaciers, there are other ways in which the visualization of rock glacier dynamics can be further enhanced, which also requires further research. For example, interactivity, in terms of letting the viewer change the viewpoint and angle, seems to be a promising feature that could enhance perception.

All in all, the addition of a diverging color scheme to a three-dimensional rock glacier animation has the potential to enhance the perception of rock glacier dynamics even for people without scientific knowledge about rock glaciers. However, there is still further research needed in order to optimize the visualization of rock glacier dynamics with and without a diverging color scheme.

References

Abermann, J., Fischer, A., Lambrecht, A., & Geist, T. (2010). On the potential of very high-resolution repeat DEMs in glacial and periglacial environments. *The Cryosphere*, 4, 53–65. www.the-cryosphere.net/4/53/2010/

Alam, F., Salih, A. E., Elsherif, M., Yetisen, A. K., & Butt, H. (2022). 3D printed contact lenses for the management of color blindness. *Additive Manufacturing*, 49, 102464. https://doi.org/10.1016/J.ADDMA.2021.102464

Ali, S., & Motala, S. (2018). GIS time series mapping of the Ciskei homeland. *South African Journal of Geomatics*, 7(3), 216–221. https://doi.org/10.4314/sajg.v7i3.1

Barbasiewicz, A., Widerski, T., & Daliga, K. (2018). The analysis of the accuracy of spatial models using photogrammetric software: Agisoft Photoscan and Pix4D. *E3S Web of Conferences*, 26. https://doi.org/10.1051/e3sconf/20182600012

Bertin, J. (1983). *Semiology of Graphics*. University of Wisconsin Press.

Bhattacharya, A., Bolch, T., Mukherjee, K., King, O., Menounos, B., Kapitsa, V., Neckel, N., Yang, W., & Yao, T. (2021). High Mountain Asian glacier response to climate revealed by multi-temporal satellite observations since the 1960s. *Nature Communications*, 12, 4133. https://doi.org/10.1038/s41467-021-24180-y

Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., Allard, M., Boike, J., Cable, W. L., Christiansen, H. H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., … Lantuit, H. (2019). Permafrost is warming at a global scale. *Nature Communications*, 10, 264. https://doi.org/10.1038/S41467-018-08240-4

Borland, D., & Taylor II, R. M. (2007). Rainbow Color Map (Still) Considered Harmful. *IEEE Computer Graphics and Applications*, 27(2), 14–17. https://doi.org/10.1109/MCG.2007.323435

Brewer, C. A., MacEachren, A. M., Pickle, L. W., & Herrmann, D. (1997). Mapping mortality: Evaluating color schemes for choropleth maps. *Annals of the Association of American Geographers*, 87(3), 411–438. https://doi.org/10.1111/1467-8306.00061

Chang, H. S., Kang, S. C., & Chen, P. H. (2009). Systematic procedure of determining an ideal color scheme on 4D models. *Advanced Engineering Informatics*, 23(4), 463–473. https://doi.org/10.1016/j.aei.2009.05.002

Chang, H. S., Kang, S. C., & Hsieh, S. H. (2007). Color schemes in 4D construction management tools. *Congress on Computing in Civil Engineering, Proceedings*, 785–792. https://doi.org/10.1061/40937(261)93

Chen, Y.-H., Tsai, M.-H., Kang, S.-C., & Liu, C.-W. (2013). Selection and Evaluation of Color Scheme for 4D Construction Models. *Electronic Journal of Information Technology in Construction*, 18. https://doi.org/10.1093/acprof:oso/9780199676200.003.0001

Cheng, S., Xu, W., Mueller, K., & Member, S. (2019). ColorMap: A Data-Driven Approach and Tool for Mapping Multivariate Data to Color. *IEEE Transactions on Visualization and Computer Graphics*, 25(2). https://doi.org/10.1109/TVCG.2018.2808489

Crameri, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. *Nature Communications*, 11, 5444. https://doi.org/10.1038/s41467-020-19160-7

Cusicanqui, D., Rabatel, A., Vincent, C., Bodin, X., Thibert, E., & Francou, B. (2021). Interpretation of Volume and Flux Changes of the Laurichard Rock Glacier Between 1952 and 2019, French Alps. *Journal of Geophysical Research: Earth Surface*, 126(9). https://doi.org/10.1029/2021JF006161

Dall'Acqua, L., Coltekin, A., & Noetzli, J. (2013, March). A comparative user evaluation of six alternative permafrost visualizations for reading and interpreting temperature information. *GeoViz_Hamburg 2013*. https://doi.org/10.5167/uzh-87817

Dawood, N., & Motala, S. (2015). Evaluating an animated and static time series map of District Six: A visual and cognitive approach. *South African Journal of Geomatics*, 4(3). https://doi.org/10.4314/sajg.v4i3.1

Diamond, E., Bernauer, T., & Mayer, F. (2020). Does providing scientific information affect climate change and GMO policy preferences of the mass public? Insights from survey experiments in Germany and the United States. *Environmental Politics*, 29(7), 1199–1218. https://doi.org/10.1080/09644016.2020.1740547

Dimri, A. P., Allen, S., Huggel, C., Mal, S., Ballesteros-Cánovas, J. A., Rohrer, M., Shukla, A., Tiwari, P., Maharana, P., Bolch, T., Thayyen, R. J., Stoffel, M., & Pandey, A. (2021). Climate change, cryosphere and impacts in the Indian Himalayan Region. *Current Science*, 120(5), 774–790. https://doi.org/10.18520/cs/v120/i5/774-790

Dübel, S., Röhlig, M., Schumann, H., & Trapp, M. (2014). 2D and 3D Presentation of Spatial Data: A Systematic Review. *2014 IEEE VIS International Workshop on 3DVis (3DVis)*. https://doi.org/10.1109/3DVis.2014.7160094

Esri. (2016). *How To: Rescale raster data*. https://support.esri.com/en/technical-article/000008671

Evangelidis, K., Papadopoulos, T., Papatheodorou, K., Mastorokostas, P., & Hilas, C. (2018). 3D geospatial visualizations: Animation and motion effects on spatial objects. *Computers and Geosciences*, 111, 200–212. https://doi.org/10.1016/J.CAGEO.2017.11.007

Feng, M., Zheng, J., Ren, J., Hussain, A., Li, X., Xi, Y., & Liu, Q. (2019). Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data. *IEEE Access*, 7, 106111–106123. https://doi.org/10.1109/ACCESS.2019.2930410

Feurer, D., & Vinatier, F. (2018). Joining multi-epoch archival aerial images in a single SfM block allows 3-D change detection with almost exclusively image information. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146(October), 495–506. https://doi.org/10.1016/j.isprsjprs.2018.10.016

Flandorfer, P. (2012). Population Ageing and Socially Assistive Robots for Elderly Persons: The Importance of Sociodemographic Factors for User Acceptance. *International Journal of Population Research*, 2012. https://doi.org/10.1155/2012/829835

Frauenfelder, R., & Kääb, A. (2000). Towards a palaeoclimatic model of rock-glacier formation in the Swiss Alps. *Annals of Glaciology*, 31, 281–286. https://doi.org/10.3189/172756400781820264

Grainger, S., Mao, F., & Buytaert, W. (2016). Environmental data visualisation for non-scientific contexts: Literature review and design framework. In *Environmental*

Modelling and Software (Vol. 85, pp. 299–318). Elsevier Ltd. https://doi.org/10.1016/j.envsoft.2016.09.004

Hardin, C. L., & Maffi, L. (Eds.). (1997). *Color Categories in Thought and Language*. Cambridge University Press. https://doi.org/10.1017/CBO9780511519819

Heitz, R. P. (2014). The speed-accuracy tradeoff: history, physiology, methodology, and behavior. *Frotniers in Neuroscience*, 8, 150. https://doi.org/10.3389/fnins.2014.00150

Hermann, M., & Carpentier III, E. (2006). Cartographic Design on Maine's Appalachian Trail. *Cartographic Perspectives*, 53, 51–56. https://doi.org/10.14714/CP53.362

Hoarau, C. (2011). Cartography and Geographic Information Science Reaching a Compromise between Contextual Constraints and Cartographic Rules: Application to Sustainable Maps. *Cartography and Geographic Information Science*, 38(2), 79–88. https://doi.org/10.1559/1523040638279

Hoarau, C., & Christophe, S. (2017). Cartographic continuum rendering based on color and texture interpolation to enhance photo-realism perception. *ISPRS Journal of Photogrammetry and Remote Sensing*, 127, 27–38. https://doi.org/10.1016/j.isprsjprs.2016.09.012

Hochschild, V., Braun, A., Sommer, C., Warth, G., & Omran, A. (2020). Visualizing Landscapes by Geospatial Techniques. In *Modern approaches to the visualization of landscapes* (pp. 47–78). Springer VS. https://doi.org/10.1007/978-3-658-30956-5_4

Immerzeel, W. W., Kraaijenbrink, P. D. A., Shea, J. M., Shrestha, A. B., Pellicciotti, F., Bierkens, M. F. P., & de Jong, S. M. (2014). High-resolution monitoring of Himalayan glacier dynamics using unmanned aerial vehicles. *Remote Sensing of Environment*, 150, 93–103. https://doi.org/10.1016/J.RSE.2014.04.025

James, M. R., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. *Journal of Geophysical Research: Earth Surface*, 117, F03017. https://doi.org/10.1029/2011JF002289

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert Scale: Explored and Explained. *British Journal of Applied Science and Technology*, 7(4), 396–403. https://doi.org/10.9734/BJAST/2015/14975

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

Koriat, A., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 107–118. https://doi.org/10.1037/0278-7393.6.2.107

Krygier, J., & Wood, D. (2005). *Making Maps: A Visual Guide to Map Design for GIS*. The Guilford Press.

Kulesza, J. A., Spencer, J. B., & Sood, A. (2017). *White Paper Standardization of Color Palettes for Scientific Visualization*. https://doi.org/10.2172/1363736

Lange, P. A. M. van, & Rand, D. G. (2021). Human Cooperation and the Crises of Climate Change, COVID-19, and Misinformation. *Annual Review of Psychology*, 73, 379–402. https://doi.org/10.1146/annurev-psych-020821

Langhammer, J., & Vackova, T. (2018). Detection and Mapping of the Geomorphic Effects of Flooding Using UAV Photogrammetry. *Pure and Applied Geophysics*, *175*, 3224–3245. https://doi.org/10.1007/s00024-018-1874-1

Levkowitz, H., & Herman, G. T. (1992). Color Scales for Image Data. *IEEE Computer Graphics and Applications*, 12(1), 72–80. https://doi.org/10.1109/38.135886

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. https://doi.org/10.1016/J.JESP.2013.03.013

MacDonald, L. W. (1999). Using color effectively in computer graphics. *IEEE Computer Graphics and Applications*, 19(4), 20–35. https://doi.org/10.1109/38.773961

Martin, D. W. (2008). *Doing Psychology Experiments* (7th Edition). Wadworth Cengage Learning.

Martínez-Carricondo, P., Agüera-Vega, F., Carvajal-Ramírez, F., Mesas-Carrascosa, F. J., García-Ferrer, A., & Pérez-Porras, F. J. (2018). Assessment of UAV-photogrammetric mapping accuracy based on variation of ground control points. *International Journal of Applied Earth Observation and Geoinformation*, 72, 1–10. https://doi.org/10.1016/j.jag.2018.05.015

Mercier, D. (2021). Climate change and the melting cryosphere. In *Spatial Impacts of Climate Change* (pp. 21–41). Wiley.

Mitasova, H., Harmon, R. S., Weaver, K. J., Lyons, N. J., & Overton, M. F. (2012). Scientific visualization of landscapes and landforms. *Geomorphology*, 137(1), 122–137. https://doi.org/10.1016/J.GEOMORPH.2010.09.033

Moreland, K. (2009). Diverging Color Maps for Scientific Visualization. In R. and P. B. and K. D. and K. Y. and W. J. and P. R. and L. P. and H. A. and E. M. L. and S. C. T. and C. D. Bebis George and Boyle (Ed.), *Advances in Visual Computing* (pp. 92–103). Springer Berlin Heidelberg.

Moreland, K. (2016). Why we use bad color maps and what you can do about it. *Human Vision and Electronic Imaging 2016, HVEI 2016*, 262–267. https://doi.org/10.2352/ISSN.2470-1173.2016.16HVEI-133

Opach, T., Midtbø, T., & Nossum, A. (2011). A NEW CONCEPT OF MULTI-SCENARIO, MULTI-COMPONENT ANIMATED MAPS FOR THE VISUALIZATION OF SPATIO-TEMPORAL LANDSCAPE EVOLUTION. *Miscellanea Geographica*, 15, 215–229. https://doi.org/10.2478/v10288-012-0013-6

Paul, F. (2015). Revealing glacier flow and surge dynamics from animated satellite image sequences: Examples from the Karakoram. *Cryosphere*, 9(6), 2201–2214. https://doi.org/10.5194/tc-9-2201-2015

Philips, A., Walz, A., Bergner, A., Graeff, T., Heistermann, M., Kienzler, S., Korup, O., Lipp, T., Schwanghart, W., & Zeilinger, G. (2015). Immersive 3D geovisualization in higher education. *Journal of Geography in Higher Education*, 39(3), 437–449. https://doi.org/10.1080/03098265.2015.1066314

Rastner, P., Joerg, P. C., Huss, M., & Zemp, M. (2016). Historical analysis and visualization of the retreat of Findelengletscher, Switzerland, 1859–2010. *Global and Planetary Change*, 145, 67–77. https://doi.org/10.1016/J.GLOPLACHA.2016.07.005

Reda, K., & Szafir, D. A. (2021). Rainbows Revisited: Modeling Effective Colormap Design for Graphical Inference. *IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS*, 27(2), 1032–1042. https://doi.org/10.1109/TVCG.2020.3030439

Roth, R. E. (2017). Visual Variables. In *International Encyclopedia of Geography: People, the Earth, Environment and Technology* (pp. 1–11). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118786352.wbieg0761

Samsel, F., Klaassen, S., & Rogers, D. H. (2018). ColorMoves: Real-Time Interactive Colormap Construction for Scientific Visualization. *IEEE Computer Graphics and Applications*, 38(1), 20–29. https://doi.org/10.1109/MCG.2018.011461525

Savage, D. M., Wiebe, E. N., & Devine, H. (2004). Performance of 2D versus 3D Topographic Representations for Different Task Types. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48, 1793–1797. https://doi.org/10.1177/154193120404801601

Seijmonsbergen, A. C. (2013). The Modern Geomorphological Map. *Treatise on Geomorphology*, 14, 35–52. https://doi.org/10.1016/B978-0-12-374739-6.00371-7

Silva, S., Sousa Santos, B., & Madeira, J. (2011). Using color in visualization: A survey. *Computers and Graphics (Pergamon)*, 35(2), 320–333. https://doi.org/10.1016/J.CAG.2010.11.015

Smart, S., Wu, K., & Szafir, D. A. (2020). Color Crafting: Automating the Construction of Designer Quality Color Ramps. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 1215–1225. https://doi.org/10.1109/TVCG.2019.2934284

Smith, M. J., Hillier, J. K., Otto, J. C., & Geilhausen, M. (2013). 3.11 Geovisualization. In J. F. Shoreder (Ed.), *Treatise on Geomorphology* (Vol. 3, pp. 299–325). Academic Press. https://doi.org/10.1016/B978-0-12-374739-6.00054-3

Steinemann, O., Reitner, J. M., Ivy-Ochs, S., Christl, M., & Synal, H. A. (2020). Tracking rockglacier evolution in the Eastern Alps from the Lateglacial to the early Holocene. *Quaternary Science Reviews*, 241, 106424. https://doi.org/10.1016/j.quascirev.2020.106424

Strozzi, T., Caduff, R., Jones, N., Barboux, C., Delaloye, R., Bodin, X., Kääb, A., Mätzler, E., & Schrott, L. (2020). remote sensing Monitoring Rock Glacier Kinematics with

Satellite Synthetic Aperture Radar. *Remote Sensing*, 12, 559. https://doi.org/10.3390/rs12030559

swisstopo. (n.d.). *swissALTI3D*. Retrieved December 4, 2021, from https://www.swisstopo.admin.ch/en/geodata/height/alti3d.html

Szafir, D. A. (2018). Modeling Color Difference for Visualization Design; Modeling Color Difference for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics*, 24(1). https://doi.org/10.1109/TVCG.2017.2744359

Tateosian, L., Mitasova, H., Thakur, S., Hardin, E., Russ, E., & Blundell, B. (2014). Visualizations of coastal terrain time series. *Article Information Visualization*, 13(3), 266–282. https://doi.org/10.1177/1473871613487086

Tucker, G. E., & Hancock, G. R. (2010). State of Science Modelling landscape evolution. *Earth Surface Processes and Landforms*, 35, 28–50. https://doi.org/10.1002/esp.1952

Tyner, J. A. (2010). *Principles of Map Design*. A Division of Guilford Publications, Inc.

Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character. Royal Society (Great Britain)*, 203(1153), 405–426. https://doi.org/10.1098/RSPB.1979.0006

Ulrich, V., Williams, J. G., Zahs, V., Anders, K., Hecht, S., & Höfle, B. (2021). Measurement of rock glacier surface change over different timescales using terrestrial laser scanning point clouds. *Earth Surface Dynamics*, 9(1), 19–28. https://doi.org/10.5194/esurf-9-19-2021

Vicentiy, A. v., Shishaev, M. G., & Oleynik, A. G. (2016). Dynamic cognitive geovisualization for information support of decision-making in the regional system of radiological monitoring, control and forecasting. *Advances in Intelligent Systems and Computing*, 466, 483–495. https://doi.org/10.1007/978-3-319-33389-2_46

Werbrouck, I., Antrop, M., van Eetvelde, V., Stal, C., de Maeyer, P., Bats, M., Bourgeois, J., Court-Picon, M., Crombé, P., de Reu, J., de Smedt, P., Finke, P. A., van Meirvenne, M., Verniers, J., & Zwertvaegher, A. (2011). Digital Elevation Model generation for historical landscape analysis based on LiDAR data, a case study in

Flanders (Belgium). *Expert Systems with Applications*, 38(7), 8178–8185. https://doi.org/10.1016/J.ESWA.2010.12.162

Wheate, R. (2012). Cartographic Depiction and Visualization of Alpine Glacier Changes in Western Canada. In New Zealand Cartography Society (Ed.), *Mountain Cartography Workshop in Taurewa* (pp. 15–25). https://doi.org/10.1.1.653.2240

Wiesmann, S., Kääb, A., & Hurni, L. (2009). VISUALIZATION OF GLACIER SURFACE MOVEMENTS. In International Cartographic Association (ICA) (Ed.), *24th International Cartographic Conference 2009*. International Cartographic Association (ICA).

Wilkening, J., Han, K., & Jahnke, M. (2019). Creating and evaluating web-based visualizations of multi-dimensional spatiotemporal data. *Proceedings International Cartography Association*, 2, 145. https://doi.org/10.5194/ica-proc-2-145-2019

Wissen, U., Schroth, O., Lange, E., & Schmid, W. A. (2008). Approaches to integrating indicators into 3D landscape visualisations and their benefits for participative planning situations. *Journal of Environmental Management*, 89, 184–196. https://doi.org/10.1016/j.jenvman.2007.01.062

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Publishing Group*, *1*, 58. https://doi.org/10.1038/s41562-017-0058

Wood, E., Klausz, N., & Macneil, S. (2021). Examining the Influence of Multiple-Choice Test Formats on Student Performance. *Innovative Higher Education*. https://doi.org/10.1007/s10755-021-09581-7

Zahs, V., Hämmerle, M., Anders, K., Hecht, S., Sailer, R., Rutzinger, | Martin, Jack, W. G., & Höfle, B. (2019). *Multi-temporal 3D point cloud-based quantification and analysis of geomorphological activity at an alpine rock glacier using airborne and terrestrial LiDAR*. https://doi.org/10.1002/ppp.2004

Zhou, L., & Hansen, C. D. (2016). A Survey of Colormaps in Visualization; A Survey of Colormaps in Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 22. https://doi.org/10.1109/TVCG.2015.2489649

Appendix

A – List of Aerial Images from swisstopo

This is the complete list of all the 287 aerial images from swisstopo being used in this thesis. "BW" stands for black and white images and "C" for color images. The number in the list is the official picture number in the swisstopo database.

1939 BW:
19390740040091
19390740010089
19390740020086
19390740030092
19390740010088
19390740030093
19390740020087
19390740040090

1946 BW:
19460750030033
19460880040041
19460820020028
19461620040081
19460850010017
19460860020018
19460770020019
19460820020032
19461620040082
19460770020020
19460820020031
19460720060039
19460830010003
19460830010004
19460870030034
19460880040042
19461620040084
19461450020043
19460820020029
19461620040083
19461450020044
19460720060040
19460750030034
19460820020030
19460830010002
19460750030035
19460830010001

1959 BW:
19599990602573
19599990602574
19599990644730
19599990602571
19599990602572
19599990602569
19599990602575
19599990602570
19599990622538
19599990612525
19599990612529
19599990612530
19599990612528
19599990644735
19599990622546
19599990612526
19599990622543
19599990644737
19599990644733
19599990622539
19599990612532
19599990644731
19599990612527
19599990622542
19599990644732
19599990622540

19599990644734
19599990622545
19599990622544
19599990622541
19599990644736
19599990612531

1961 BW:

19619990460239
19619990490309
19619990480570
19619990460238
19619990410281
19619990460245
19619990460244
19619990460243
19619990470337
19619990460241
19619990460240
19619990460242
19619990480571
19619990490308
19619990470332
19619990480572
19619990490304
19619990480576
19619990470331
19619990470336
19619990490306
19619990480573
19619990470334
19619990480575
19619990490305
19619990490307
19619990470333
19619990470335
19619990490303
19619990480577
19619990480574
19619990470330

1973 BW:

19739990479659
19739990479658
19739990479657
19739990479656
19739990479661
19739990479660
19731770049730
19731800041333
19731770049725
19731800041334
19731800041332
19731770049727
19731800041330
19731800041336
19739990489654
19731770049724
19731770049729
19731800041331
19731770049728
19731770049726
19731800041335

1978 BW:

19789990263354
19781770023535
19781770023541
19781800023490
19781800023488
19781800123491
19781800023489
19781800023486
19781770023540
19781800023487
19781770023538
19781770023536
19781770023537
19781800123492
19781770023539

1985 BW:

19851760054394
19851800044369
19851760054393
19851800044370
19851800044373
19851760054396
19851760054397
19851760054399
19851800044371
19851800044372
19851760054395
19851800044375
19851800044374
19851760054398

1991 BW:

19911800027527
19911760017187
19911760017186
19911800027526
19911800027523
19911800027524
19911760017183
19911760017182
19911800027525
19911760017185
19919990067075
19911760017184
19919990067074
19919990046362
19911800027522

1997 BW:

19971800053681
19971760048584
19971760048585
19971800053680
19971800053679
19971800053677
19971760048588
19971760048586
19971800053675
19971760048590
19971800053678
19971760048587
19971800053676
19971760048589

1999 BW:

19990050145275
19990050145277
19990050135294
19990050145276
19990050135295
19990050135290
19990050135293
19990050145278
19990050145282
19990050145279
19990050145284
19990050135292
19990050145280
19990050135291
19990050145281
19990050145283
19990050135289

1999 C:

19991991742274
19991991512249
19991991512251
19991991732285
19991991742272
19991991512257
19991991742264
19991991732277
19991991732284
19991991742265
19991991742266
19991991512254
19991991512253
19991991742273
19991991742270
19991991742269
19991991732276
19991991512250
19991991512256
19991991732280
19991991742267
19991991512252
19991991742271
19991991512255
19991991732282
19991991732283
19991991732278
19991991732279
19991991732281
19991991742268

2000 C:

20001994741066
20001994511000
20001994741065
20001994731067
20001994741063
20001994511002
20001994511008
20001994511001
20001994731069
20001994511004

20001994741064
20001994731071
20001994731072
20001994741061
20001994731068
20001994741060
20001994741057
20001994741058
20001994511007
20001994511003
20001994741062
20001994731074
20001994731075
20001994511005
20001994731070
20001994731073
20001994741059
20001994511006

2003 C:

20031811040446
20031811040447
20031841070413
20031751043459
20031781040486
20031751043458
20031751043460
20031781040487
20031841070415
20031811040444
20031841070417
20031841070414
20031841070416
20031751043462
20031811040445
20031781040488
20031751043461
20031781040490
20031781040489
20031811040443

B – Steps for Smoothening

To smoothen the DEMs to get rid of the bumps and holes in them, the following steps in ArcGIS Pro were necessary:

1. Slope tool: Create slope layer from the DEM layer
2. Reclassify tool: reclassify slope > 55 = 1; rest of the values = NoData
3. Raster to polygon tool: run function, make sure that the box “simplify” is unchecked
4. Buffer tool: Select a buffer of 1m (why 1m is explained in section 3.2.2).
5. Polygon to Raster tool: Select cell size equal to the original DEM raster
6. Raster calculator: Con(IsNull('Raster from step 5'), 'Raster DEM original')
7. Raster to points tool
8. Create TIN tool
9. TIN to raster tool
10. Clip tool: clip new DEM with the outlines of the rock glacier
11. Hillshade tool

The reclassification of the slope allowed to extract all values with a steep slope. The reason why this was done is that the bumps and holes appeared mainly in steep areas. After this step the DEM, and consequently the polygon layer, was holey. Therefore, steps 7-9 were necessary to make the layer continuous again.

C – Links to Animations

Links for the final animations:

Animation with color: https://www.youtube.com/watch?v=3xyrQ7jtuM0

Animation without color: https://www.youtube.com/watch?v=Gs6PFgrt2yU

D – Survey

D.1 – Survey Structure

The Representation of Rock Glacier Dynamics in 3D Animations

This survey is part of my master's thesis at the University of Zurich. I am investigating the representation of rock glacier dynamcis in 3D animations. As a participant, you will see an animation of a rock glacier on the Macun Plateau in the Swiss National Park for a timespan from 1959 to 2003.

The survey consists of 10 question groups mainly regarding the rock glacier dynamics, focusing on the change of volume. On each page you will be able to watch the animation again, as many times as you want. Note that there is a question at the end of each page asking how many times you have watched the animation **on this page**. Please try to count correctly.

This survey involves the collection of your personal data. All data will be anonymised. Furthermore, your name will not be used in the work. By clicking on the "Next"-button at the lower right corner of this page you acknowledge this and further agree that your data can be published in anonymised form for scientific purposes.

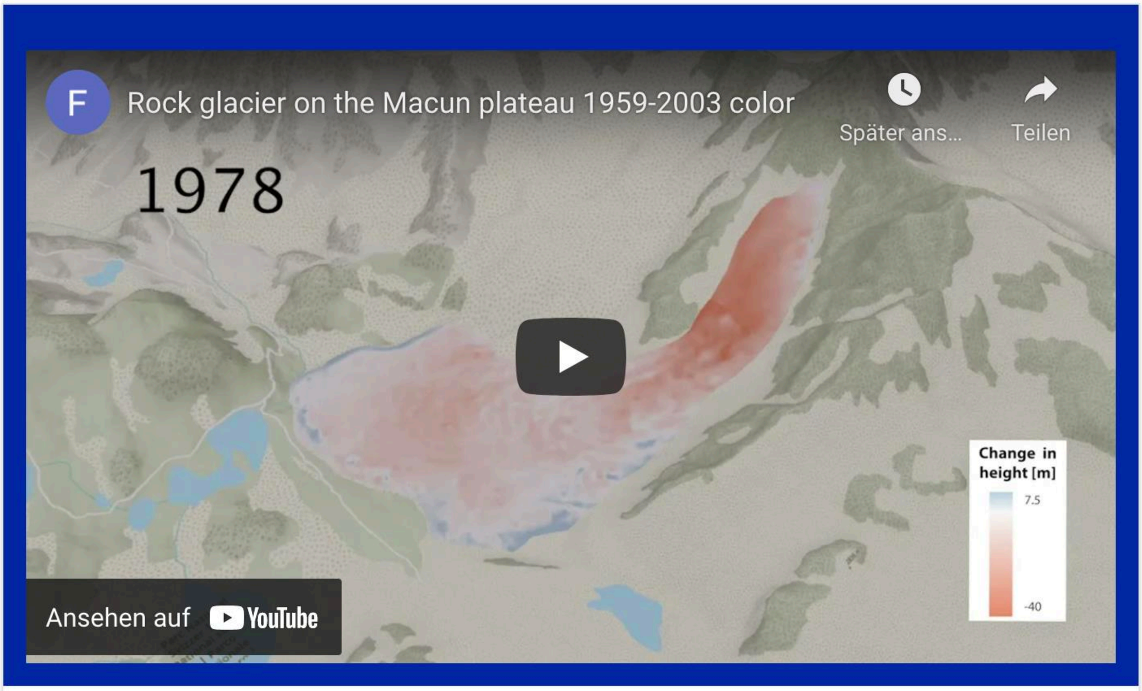
If you have any questions or remarks about the survey, do not hesitate to contact me (flo.flueckiger@uzh.ch).

I really appreciate your input!

There are 54 questions in this survey.

Next

Question 1



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.

* Did the volume of the rock glacier increase or decrease over the whole area in the timespan from 1959 to 2003?

Choose one of the following answers

Increase Decrease

* How confident are you with your answer?

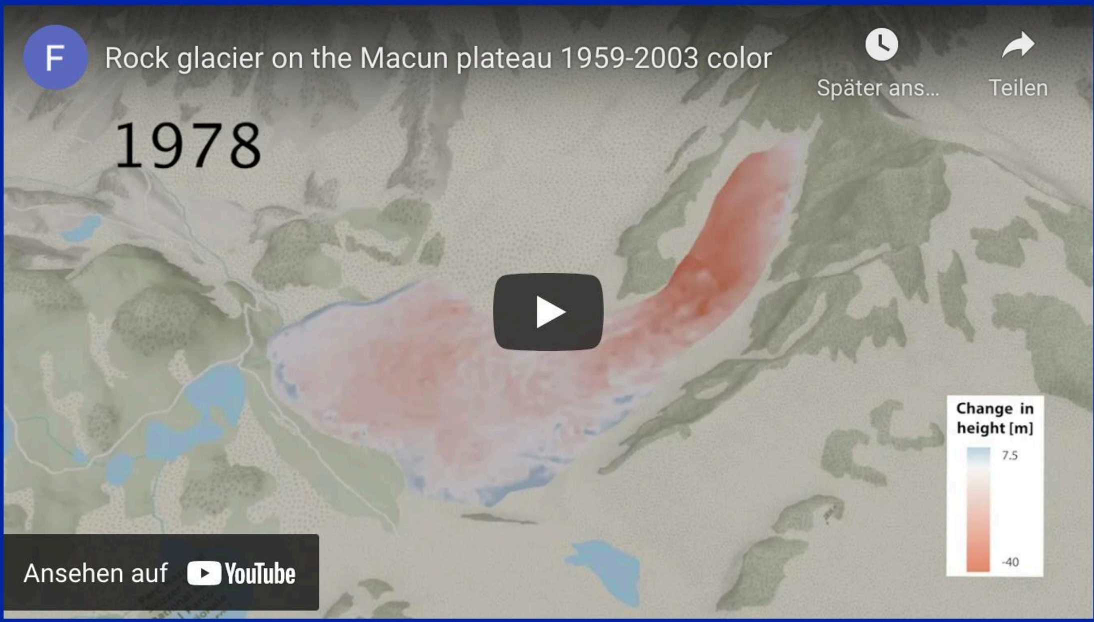
○ 1 ○ 2 ○ 3 ○ 4 ○ 5

1 = not confident, 5 = very confident

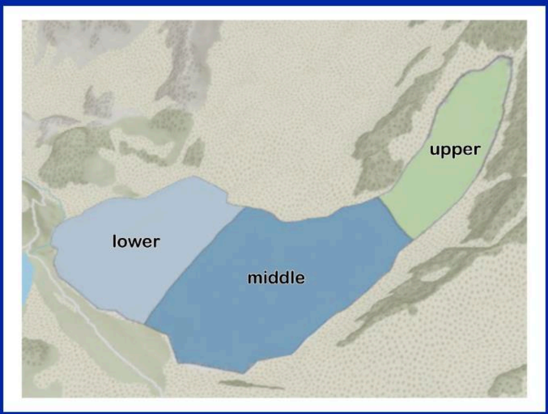
* How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 2



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.



To clarify what is meant by lower, middle, and upper part

* In which part of the rock glacier (upper, middle, lower) did the volume of the rock glacier change the **least** in the timespan from 1959 to 2003?

Choose one of the following answers

Upper Middle Lower

* In which part of the rock glacier (upper, middle, lower) did the volume of the rock glacier change the **most** in the timespan from 1959 to 2003?

Choose one of the following answers

Upper Middle Lower

* How confident are you with your answer?

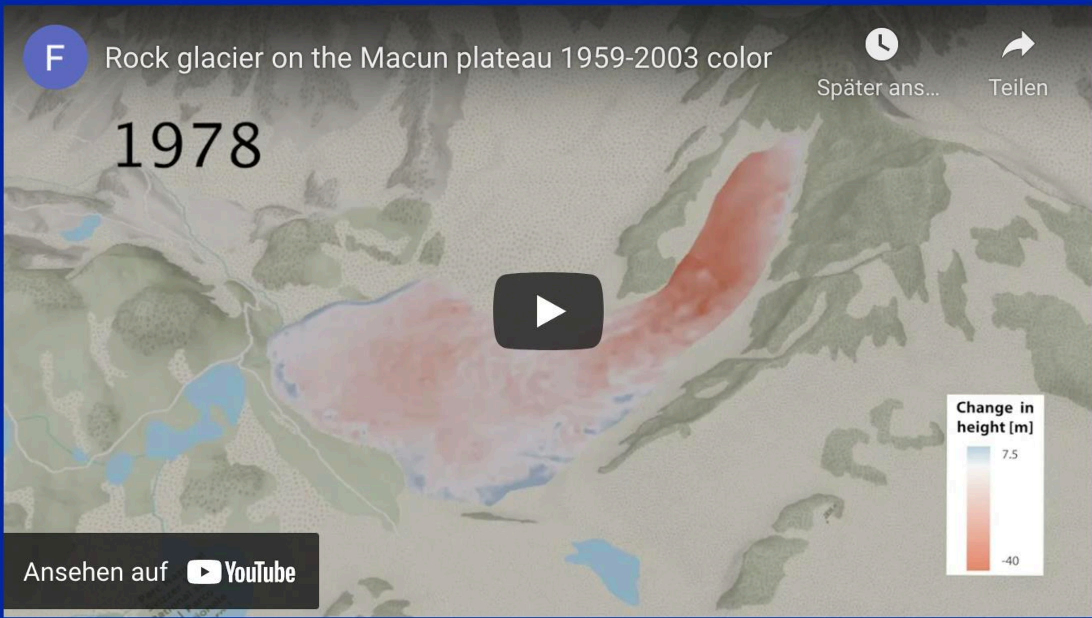
○ 1 ○ 2 ○ 3 ○ 4 ○ 5

1 = not confident, 5 = very confident

* How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 3



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.

* Was the **decrease** in volume over the whole area stronger in the period from 1959-1969 or 1978-1988?

Choose one of the following answers

1959-1969

1978-1988

✱ How confident are you with your answer?

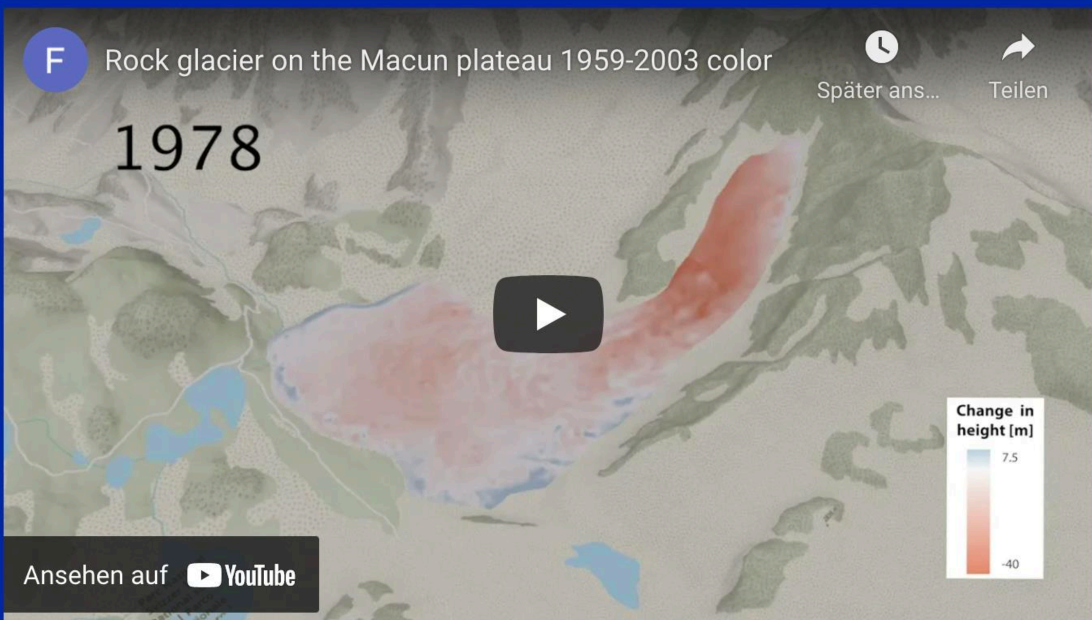
◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

1 = not confident, 5 = very confident

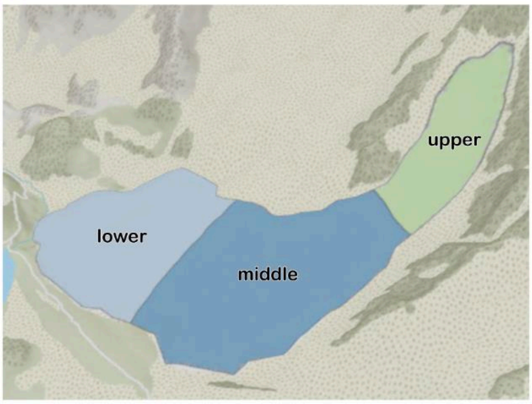
✱ How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 4



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.



To clarify what is meant by lower, middle, and upper part

* Was the **decrease** in volume in the **lower** part of the rock glacier stronger in the period from 1969-1973 or 1988-1993?

Choose one of the following answers

1969-1973 | 1988-1993

* How confident are you with your answer?

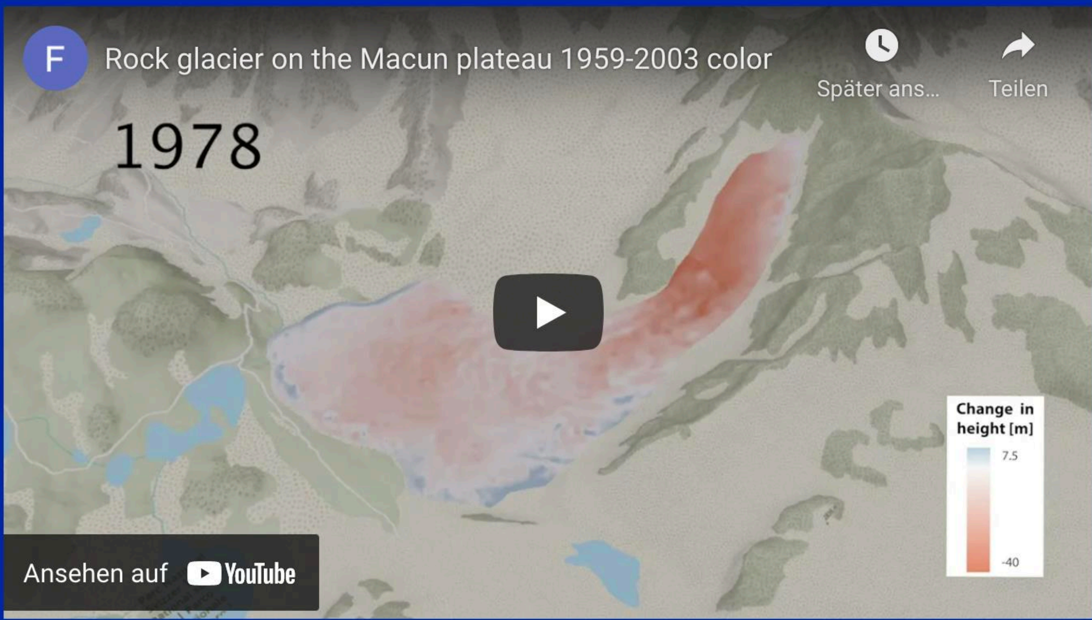
○ 1 ○ 2 ○ 3 ○ 4 ○ 5

1 = not confident, 5 = very confident

* How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 5



❓ The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.

* Did the volume of the rock glacier over its whole area decrease **faster** (rate of change in meters per year) before or after 1983?

❗ Choose one of the following answers

Before   After

* How confident are you with your answer?

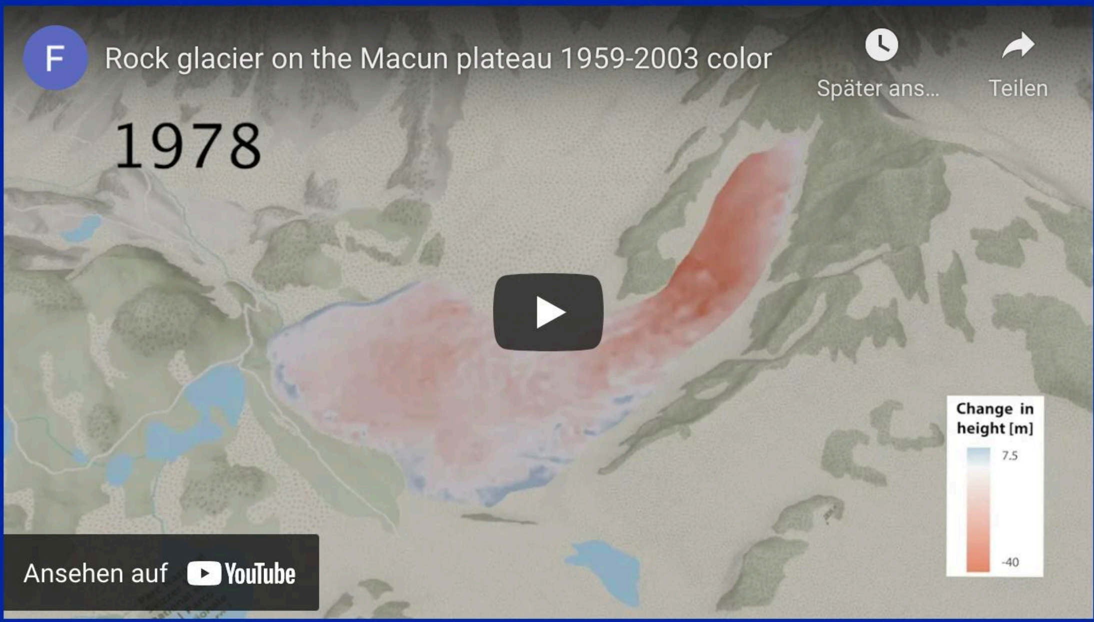
◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

❓ 1 = not confident, 5 = very confident

* How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 6



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.

* How much did the volume of the rock glacier over the whole area change in the timespan from 1959 to 2003? Make an estimation in %. Please indicate "-" for decrease and "+" for increase.

Only numbers may be entered in this field.

* How confident are you with your answer?

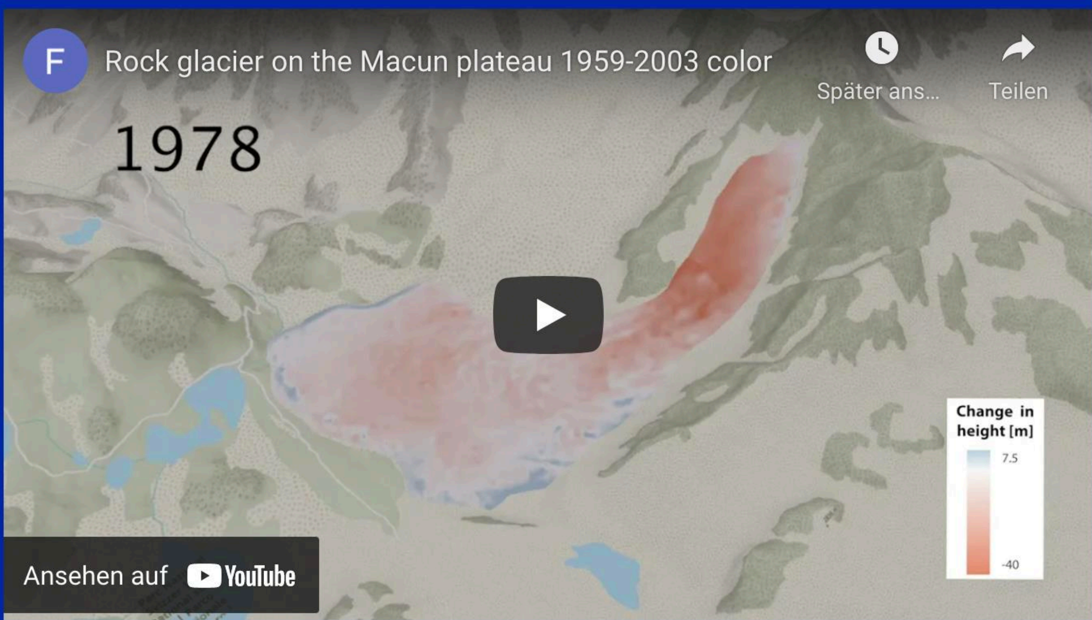
○ 1 ○ 2 ○ 3 ○ 4 ○ 5

1 = not confident, 5 = very confident

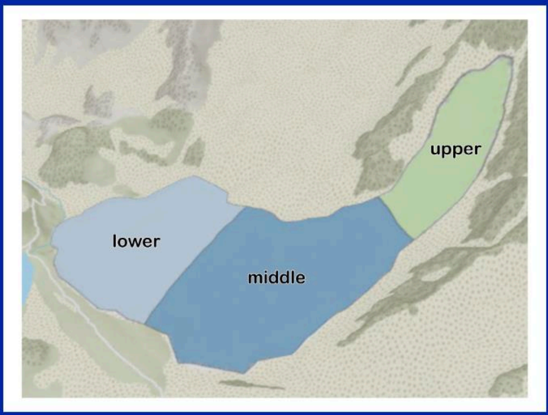
* How many times did you watch the animation on this page?

Only numbers may be entered in this field.

Question 7



The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.



❓ To clarify what is meant by lower, middle, and upper part

* How much did the volume in the **upper** part of the rock glacier change in the period from 1973-1978? Make an estimation in %. Please indicate "-" for decrease and "+" for increase.

❗ Only numbers may be entered in this field.

* How confident are you with your answer?

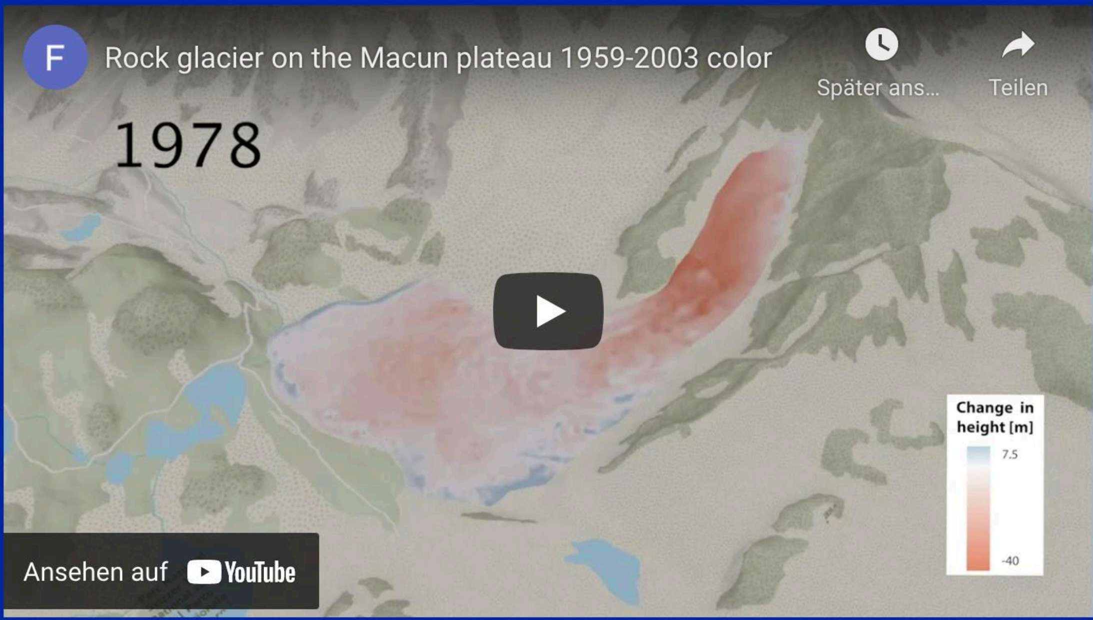
○ 1 ○ 2 ○ 3 ○ 4 ○ 5

❓ 1 = not confident, 5 = very confident

* How many times did you watch the animation on this page?

❗ Only numbers may be entered in this field.

Question 8



❓ The color scheme shows the difference in height in meters. A change in height also indicates a change in volume.

What else did you observe besides the changes in volume?

* How many times did you watch the animation on this page?

❗ Only numbers may be entered in this field.

Sociodemographic Questions

*** How old are you?**

Only numbers may be entered in this field.

*** What gender do you identify as?**

Choose one of the following answers

Male Female Other Prefer not

*** What is your field of study or work?**

*** How familiar are you with geomorphological processes, such as rock glacier dynamics?**

○ 1 ○ 2 ○ 3 ○ 4 ○ 5

1 = very unfamiliar, 5 = very familiar

General questions

* How difficult was the answering of the questions for you overall?

◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5

1 = very easy, 5 = very difficult

Do you have any further comments?

You have reached the end of this survey. Thank you for your participation.

D.2 – Correct Answers of Survey

Question 1: Decrease (mean = -8.03m)

Question 2: **a.** Lower (mean upper = -17.54m, mean middle = -6.61m, mean lower = -6.14m)

b. Upper (mean upper = -17.54m, mean middle = -6.61m, mean lower = -6.14m)

Question 3: 1959-1969 (mean 1959-1969 = -4.85m, mean 1978-1988 = -0.26m)

Question 4: 1969-1973 (mean 1969-1973 = -1.46m, mean 1988-1993 = -0.09m)

Question 5: Before (1959-1983 = -0.22 m/a, 1983-2003 = -0.13 m/a)

Question 6: -0.30%

Question 7: +0.16%

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.

Flo Flückiger

Herrliberg, 15.09.2022