Designing for Ambiguity in Sensemaking: Visual Analytics in Risk Analysis and Prediction

by

Stan Nowak

B.A., University of British Columbia, 2014

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
School of Interactive Arts and Technology
Faculty of Communication, Art and Technology

SIMON FRASER UNIVERSITY
Fall 2023

Declaration of Committee

Name:	**Stan Nowak**
Degree:	**Doctor of Philosophy**
Title:	**Designing for Ambiguity in Sensemaking: Visual Analytics in Risk Analysis and Prediction**
Committee:	**Chair: Kate Hennesy** Associate Professor, Interactive Arts and Technology
	Lyn Bartram Supervisor Professor, Interactive Arts and Technology
	Pascal Haegeli Committee Member Associate Professor, Resource and Environment Management
	Wolfgang Stuerzlinger Committee Member Professor, Interactive Arts and Technology
	Brian Fisher Examiner Professor, Interactive Arts and Technology
	Miriah Meyer External Examiner Professor, Science & Technology Linköping University

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Abstract

This dissertation investigates how visual analytics tools and techniques can address ambiguity in complex risk assessment, prediction, and monitoring, focusing on the domain of avalanche forecasting. Drawing on a broad set of methods and theory from complex cognitive systems engineering and visualization research, this dissertation delves into the cognitive work demanded by this domain and explores visual analytics solutions to enhance sensemaking.

In a study using a variety of methods including interviews, observational research, and situated-recall, this research identifies and characterizes the issues of ambiguity in avalanche forecasting as they pertain to individual and collaborative sensemaking around data. It presents the results of a participatory design study that develops visualization tools to tackle these challenges and an evaluation study investigating the analytic affordances and sensemaking support provided by newly designed and existing tools used by forecasters. In addition, a preliminary study using participatory design and diary study methods investigates how knowledge construction and synthesis can be supported to better address challenges of shared sensemaking in asynchronous sequential collaboration.

Findings from this dissertation reveal the shortcomings of conventional visualization guidelines in being able to tackle ambiguity in this complex domain. Instead of employing efficient and effective perceptual encodings and summary overviews, it highlights the significance of flatter visual hierarchies, visual difficulty, and rapid access to details for better support of sensemaking around ambiguity. In addition, it reveals new challenges and opportunities for improved knowledge synthesis support in visual analytics tools. The theoretical framing and methodological approach used in this dissertation is novel for the domain of visual analytics.

Keywords: Visual Analytics; Sensemaking; Complex Systems; Avalanche Forecasting; Cognitive Systems Engineering

Dedication

I dedicate this thesis to my parents and loving life partner.

Mom and dad, thank you for your support and guidance that shaped me into who I am today.

Megan, you are a constant source of inspiration. Your tenacity and ambition is infectious. Thank you so much for your loving support throughout this journey.

Acknowledgements

Thank you, Dr. Lyn Bartram, for taking the chance and believing in me. I feel so grateful to have had you as my supervisor. You always told me what I needed to hear and were an unwavering and patient supporter of my work. A true mentor.

Thank you, Dr. Pascal Haegeli, your advocacy was instrumental in allowing this research to move forward and your guidance kept me on the right track.

Dr. Wolfgang Steurzlinger, without those early conversations in your visualization course, I might have ended up doing an entirely different thesis. Thank you for your guidance and mentorship.

Thank you to all the forecasters and staff at Avalanche Canada who have been invaluable supporters, advocates, and incredible collaborators of my work. Special thanks to Dr. Simon Horton, Karl Klassen, and Grant Helgeson.

Thank you to all past and present SFU Avalanche Research Program lab-mates. I am lucky to have met and benefitted from such a generous and supportive group of researchers.

Thank you to all the SIAT faculty and staff. I absolutely loved all your courses and your dedication to graduate students is exceptional. Thank you for striving to uphold your mission as an institution. The knowledge I gleaned from your work was a constant presence throughout my PhD journey.

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

Complex analyses are characterized by constructive sensemaking processes. The inclination for humans to make connections between actors, events, or otherwise incomplete information is often framed in terms of the risks it poses for making errors in judgement (Tversky & Kahneman, 1974), particularly in context of data analysis. As much as this "magical" way of thinking might not align with standard models of rational decision making or statistically sound inferences, it is a natural part of scientific and analytic inquiry (Diaconis, 2006). Human observation is inherently limited (Grabowski & Strzalka, 2008), rendering comprehensive understanding from data alone intractable (Grabowski & Strzalka, 2008). Instead, complexity demands the constructive knowledge-based process of selecting from known explanations, generating new ones and evaluating them based on the fragmented incomplete information or data available. This is what Klein et al. (2007) refer to as sensemaking. Sensemaking is therefore much more about resolving *ambiguity* than simply seeking more information to fill gaps. This ambiguity, the state where multiple interpretations are plausible, is a functional component of sensemaking producing both costs and benefits to analytic work. Ambiguity is particularly pronounced when analytic work is constrained, involves risk-based decision-making, assessment, or prediction, and or involves collaboration (Heer & Agrawala, 2008).

Visual analytics, defined as the "science of analytical reasoning facilitated by interactive visual interfaces" is well suited to address the challenges of ambiguity (Cook & Thomas, 2005). Understanding sensemaking is considered to be a foundational aspect of how to design visual analytics systems (D. Keim et al., 2008). However, the most prominent models of sensemaking used in visual analytics (Pirolli & Card, 2005; Shrinivasan & van Wijk, 2008), developed in the context of exploratory intelligence analysis, do not fully capture the nuances of ambiguity or the challenges of risk prediction and management domains. Cognitive models, in particular, are under-utilized in guiding visualization research and design (Padilla, 2018). Visualization researchers have produced a rich literature on how visualizations and the perceptual processes involved (Cleveland & McGill, 1984) do or do not facilitate sound inferences (Bertini et al., 2020) and insights (North, 2006), particularly when dealing with data uncertainties (Kamal et al., 2021). These studies have been and will continue to be instrumental for understanding low-level cognitive and perceptual processes with great precision, control, and generalizability. As important as these considerations are in visualization design, they neglect the higher-level knowledge-based cognitive processes and richness of the real-world context in which visual analytics is applied. This is why some have argued that the goal of visual analytics is not the generation of insights but rather the iterative development and calibration of mental models to some end such as an assessment, prediction, or a decision (Andrienko et al., 2018).

Prior knowledge can directly affect how visualizations are perceived and the inferences made (Xiong et al., 2019). Many visualization researchers acknowledge the role that interpretation plays in uncertainty more generally (Boukhelifa et al., 2017; MacEachren et al., 2005; MacEachren, 2015; Thomson et al., 2005; Zuk & Carpendale, 2007). For instance, Liu et al. (2020) describe 'cognitive alternatives' as alternative hypotheses, mental models, and interpretations in visual analysis. Meanwhile, Lin et al. (2021) define the idea of a 'data hunch' as "a person's knowledge about how representative data is of a phenomenon of interest" and outline a design space for externalizing data hunches. Similarly, researchers identify 'implicit errors', errors inherent to a data set but not explicitly represented within it, as a challenge to be address with targeted visualization design (Panagiotidou et al., 2021) or targeted knowledge elicitation mechanisms (Mccurdy et al., 2019).

Making the implicit or tacit knowledge-based processes involved in analysis explicit is a common solution employed in visual analytics. This is the motivation for more recent developments in visual analytics approaches for human-machine teaming where human knowledge is embedded and modeled using algorithms towards some targeted problem solving application (Federico et al., 2017; Green & Ribarsky, 2008; Rind et al., 2019), for instance in decision-support tools for clinical gait analysis (Wagner et al., 2019). Knowledge elicitation is also common in collaborative visual analytics as way to facilitate collaborative sensemaking through communication (Goyal et al., 2014; Heer & Agrawala, 2008; P. Isenberg et al., 2011; Mathisen et al., 2019; Prue et al., 2014; Shrinivasan & van Wijk, 2008; Zhao et al., 2018).

As ambiguity is often viewed as a problematic, rather than a functional component of sensemaking, this diagnosis naturally leads to solutions that aim to reduce ambiguity. Externalizing tacit knowledge to make it explicit is an example of this. While this may be appropriate under some circumstances, ambiguity can itself serve a productive role – particularly when conveying uncertainty (Sterner, 2022) or incomplete analysis. To date, there has been little investigation into how the design of visual analytics systems can support or mediate ambiguity and the specific related sensemaking processes as it pertains to the needs of a particular context of application. This thesis aims to fill this gap by identifying and diagnosing specific instances of ambiguity and exploring targeted visual analytics solutions to support the involved sensemaking process.

Researchers have underscored that designing visual analytic tools to address high-level knowledge-based cognitive processes, such as those pertaining to ambiguity, requires understanding the reality, broader context, knowledge constructs (Andrienko et al., 2018), and cognitive processes (Hollnagel & Woods, 2005) involved within a particular domain of application. Ambiguity involves complex, context-dependant, and difficult-to-anticipate matters. It is therefore inappropriate to study in a synthetic lab-based setting. Any investigation targeting ambiguity needs to preserve the naturalistic setting in which it arises.

This dissertation is therefore conducted in collaboration with and applied to the complex collaborative sensemaking and risk-based domain of avalanche forecasting. It details a multi-year long-term collaboration with avalanche forecasters at Avalanche Canada, one of Canada's primary public avalanche warning services, to develop visual analytics systems addressing the core challenges and needs of this domain. Snow avalanches are natural disasters that pose significant risks to human life and infrastructure. They are caused by structural instabilities in the continuously evolving seasonal snowpack releasing destructive masses of snow (McClung, 2002a). Avalanche forecasters are responsible for analyzing the current state of the snowpack, relating it to numerous terrain and weather conditions, avalanche activity, and potential release triggers from natural and human interactions, to determine the likelihood and potential destructive size of potential avalanches. Due to the variability of the data, much of it human-generated, and the dynamic complexity of this natural phenomenon, avalanche forecasting comprises highly complex and uncertain analyses and prediction tasks involving heavy use of expert knowledge and experience to make sense of heterogeneous, diverse, and incomplete data from multiple sources. Forecasters also typically work in distributed teams relying on continuous monitoring and analysis that builds on itself as the snowpack evolves. These characteristics make avalanche forecasting an ideal application domain to investigate visual analytics interventions aiming to address the challenges of complex analyses in risk-based contexts.

The present dissertation aims to characterize domain problems and design, develop, and evaluate visual analytics solutions in risk assessment, risk prediction, and collaboration through application-driven research in avalanche forecasting. This dissertation is structured in the following way. In Chapter 2, I provide relevant background information on ambiguity, complex systems, sensemaking theory, and related work in visual analytics and visualization. In Chapter 3, I describe the problem domain of avalanche forecasting discussing goals, relevant constructs, reasoning processes, available data and information systems, the related known domain challenges, and additional contextual information about Avalanche Canada and our research engagement. In Chapter 4, I outline my research approach and methodologies used. Chapter 5 presents findings from a two-part study involving foundational interview, observational, and situated-recall methods characterizing the problem domain of avalanche forecasting. This surfaced the issues of ambiguity as they pertain to data, analytic processes, and collaboration. This is followed by the description of a study that used participatory design methods to develop visualization tools addressing the challenges of ambiguity in Chapter 6. Chapter 7 is dedicated to an evaluation study investigating the analytic affordances and ambiguity support capacity of the prototype developed in Chapter 6 as well as some of the existing tools used by forecasters already. In Chapter 8, I report on a preliminary three-part participatory design study utilizing diary study methods to explore potential tools and techniques for knowledge capture during analysis to address further challenges of ambiguity identified in the Chapter 5 Study. In Chapter 9, I discuss key lessons learned, how this thesis addresses the stated research objectives, the primary contributions produced, design implications to address ambiguity informed by this research, future work, and limitations.

1.1. Contributions

The results and understanding generated in this dissertation challenge traditional visualization and interaction design guidelines. Specifically, one of the key lessons drawn is that always starting with summary overviews with detailed information only available on demand, as per the information seeking mantra “Overview first, zoom and filter, then details on demand” (Shneiderman, 1996), may often be problematic and inappropriate in this application problem area. Findings challenge the convention of strict adherence to design decisions made purely based on strong visual hierarchies and the use of perceptual salience to direct attention. I instead argue for the use of flatter visual hierarchies, encouraging the viewer to gain more control of their attention and consider alternative perspectives in data.

This dissertation contributes a formative exploration of the practical implications of treating ambiguity as a fundamental aspect of sensemaking processes rather than scourge to always and unilaterally reduce or remove. The research makes the following contributions to the field of visual analytics:

a) a case study of visual analytics applied in a real-world complex risk prediction and management domain;

b) a characterization for how ambiguity arises in the domain of avalanche forecasting to inform the design of visual analytics solutions in risk assessment, prediction, monitoring, and collaboration;

c) a preliminary set of explorations of interactive visualization design strategies and the resulting guidelines for tools to support ambiguity in sensemaking;

d) A qualitative framework and method for evaluating visual analytics tools in complex systems derived from cognitive systems engineering;

e) An evaluation of how different representations can serve to enhance or impede ambiguous sensemaking; and

f) a visual analytics system designed to address ambiguity that has seen field deployment.

In addition, this doctoral research has also made several practical contributions to avalanche forecasting. These include: a) the results of cognitive task analyses detailing workflows and reasoning processes; b) publications applying visualization design principles to make snowpack simulation models more easily interpretable and more operationally viable (Horton et al., 2018, 2020); c) conference presentations on applying visual analytics to avalanche forecasting (Nowak, 2019a, 2021); d) a magazine article articulating the benefits of applying visual analytics to avalanche forecasting (Nowak, 2019b); and finally, e) a set of visual analytics tools and techniques that have been deployed for operational use and are providing value for avalanche safety organizations including but not limited to Avalanche Canada, Avalanche Quebec, Parks Canada, and the Colorado Avalanche Information Center.

Chapter 2.

Background & Related Work

My research is informed by relevant background information and related work, which I present in detail in this chapter. I first define ambiguity by distinguishing it from uncertainty and discussing the relevant sensemaking and cognitive processes involved, their relation to complexity. In addition, I draw on related information processing and cognitive theories on the use of representation for problem solving activities. Finally, I review literature related to the topic of ambiguity in visual analytics and collaborative visual analytics.

2.1. Ambiguity and Uncertainty

It is perhaps fitting that the term *ambiguity* carries different meanings in everyday speech and different communities of practice. Ambiguity is often described as a form of uncertainty associated with data or quantitative information. The concept of ambiguity aversion, for example, was famously popularized by the economist Daniel Ellsberg in thought-experiments where he conjectured people prefer choices for which probabilistic information was known rather than those in which the odds are unknown or ambiguous (also known as 'knightian uncertainty') (Ellsberg, 1961). Similarly, various taxonomies from fields such as statistics, psychology, or engineering classify ambiguity as a type of uncertainty involving multiplicities or vagaries (Ayyub, 2001; Smithson, 1989). For a thorough review of different notions of ambiguity as well as an overview of relevant considerations from the perspective of risk assessment, see the review provided by Johansen and Rausand (2015) where ambiguity is defined as "the existence of multiple interpretations concerning the basis, content, and implications of risk information".

As uncertainty draws strong connotations of uncertainty dealing with data specifically, I treat ambiguity as a distinct issue. It is a product of the epistemological stance between a human observer and the world, rather than a property of data itself. In this sense, ambiguity is more a question of human interpretation and reasoning processes. While uncertainty can be a part of what instantiates ambiguity, they are distinct concepts. In this sense, the notion of ambiguity used in this thesis is much more closely related to the semiotic or philosophical notions of ambiguity (Sennet, 2016).

2.2. Sensemaking and Ambiguity

Sensemaking, the everyday reasoning where meaning is constructed from information and experience, is brought on by uncertainty, ambiguity, and when expectations are violated (Maitlis & Christianson, 2014; Weick, 1995). It is a means to cope with the complexity of our world. Human observation is inherently limited (Grabowski & Strzalka, 2008), rendering comprehensive understanding of complex systems intractable (Kirsh, 2010). Rather than mechanistic reduction of component parts, sensemaking treats complexity holistically through the consideration of alternative explanations which cohere with observed information. Sensemaking is more than just summarizing or accounting for missing information, it is about resolving multiple potential meanings.

Understanding sensemaking is a foundational pillar of visual analytics (D. A. Keim et al., 2008), but well-known models of sensemaking in analytics, notably (D. A. Keim et al., 2008; Pirolli & Card, 2005; Shrinivasan & Van Wijk, 2008) tend to be generic and do not fully capture the nuances of reasoning in risk-based domains. For instance, the sensemaking model by Pirolli and Card (2008), developed in the context of in the context of intelligence analysis, is well-suited to explain key aspects of exploratory visual analysis involving emerging insights and discovery. By contrast, risk prediction and safety management work face additional challenges of uncertainty, time or resource constraints, ill-defined goals, distributed work roles, and decision-making from incomplete and varied data (Hollnagel & Woods, 2005; Johansen & Rausand, 2015; Smith & Hoffman, 2017). These additional demands result in specific sensemaking processes that must be considered in the design of technology used.

I draw instead from cognitive research in complex systems, particularly crisis and risk management, (Klein et al., 2006) using the model of *frames* as explanatory structures to interpret, understand, and organize data into patterns (Klein et al., 2007). Frames might include causal relationships, chronologies, or procedures that may be involved in shaping data. They fill in the gaps left by fundamentally incomplete data. Frames set expectations for what counts as data and therefore guide the search for more data. Data and frames mutually determine each other, in that, data are used as cues to identify relevant frames, and conversely, frames determine which data are noticed or sought.

When data do not fit with the expectations of a given frame, ambiguity arises and sensemaking begins. It is an active process that involves seeking further relevant data and actively improving or replacing a given frame with one that is better matching. This involves iterative cycles of mapping frames and data to each other, augmenting these frames, questioning and comparing frames, and replacing frames through reframing. Thus, sensemaking relies heavily on abductive reasoning – inferences to the best available explanation. This model of sensemaking, developed through study of crisis situations, has proven robust in explaining human sensemaking processes in a variety of applications within risk-based contexts (Klein et al., 2006).

Researchers studying sensemaking in risk prediction applications such as weather forecasting (Hoffman et al., 2017) have identified sensemaking processes with targeted functions. These are closely related to the data-frame theory of sensemaking in that they emphasize considering alternative explanations under different situational demands. Relevant to the forecasting of avalanches is *anticipatory thinking* which involves mental preparation for potential risks, many of which might be highly unlikely but could result in severe consequences (Klein & Snowden, 2011). It involves directing attention to often subtle and contextually sensitive cues that could signal threats while at the same time maintaining sensitivity to cues that deviate from expectations and challenge understanding. '*Problem detection*', when an observer becomes aware of a threat that might require a course of action, depends on the observers existing understanding to compare data against (Klein et al., 1999, 2005). Often, anticipatory thinking also involves extrapolating trends into alternative future scenarios and planning for them. This exploratory and imaginative planning activity is often referred to as *'mental simulation'* (Klein & Crandall, 2018) and is mark of competence in avalanche forecasting (Adams, 2005) and weather forecasting (Pliske et al., 2004).

2.2.1. Gisting

I draw from more general research in information processing theory (Gamino et al., 2010; Ju & You, 2018) to identify a further sensemaking process that is often overlooked in analysis but is considered an essential component of critical thinking and reasoning: *gisting*. Familiar to scholars of reading (Elfenbein, 2018), a gist is not simply an overview or a summary. Instead, gist-reasoning involves assimilating and interpreting incoming information to derive global meaning from explicit details, capturing the essence of a situation rather than a simple collection of facts (Elfenbein, 2018; Ju & You, 2018). Vision scientists describe the “gist of a scene” as the phenomenon where an observer can rapidly comprehend the meaning of a complex scene at a glance without conscious effort or attention (Rensink, 2000; Ware, 2022). The gist sets expectations for what objects might be found in a scene and in doing so facilitates the detection of distinct objects (Rensink, 2000). More recently, a study of how data workers use analytic tools found them returning to raw or base data tables (Bartram et al., 2021), using gisting to develop an understanding of the data and its meaning. These processes involved scanning or reading raw base data to see overall structure, detect problems and holes, find and validate data, and understand the types of questions the data afforded.

It is important to acknowledge that the concept of gist, its role in visual sensemaking and memory, and the notion of dual-processes in cognition more generally, has roots in the Gestalt school of psychology (Reyna, 2012). Gestalt, which emerged as a reaction to ‘structuralist’ schools of psychology, emphasizes a perspective of holism where psychological phenomena are more than their atomic constituent parts and where the whole and its relation to constituent parts better describes these phenomena (Koffka, 1935). A ‘gestalt’, loosely translated from German as ‘pattern’ or ‘configuration’, describes an experience that goes beyond the summation of constituent parts and is stable despite different configurations and presentations. Gestalt has been most influential in the study of perception, describing various organizing principles for how meaning is derived from the structure of sensory information like vision (Ware, 2019). However, the ideas of holism, simplicity, and essential meaning in Gestalt psychology have extended to higher-level cognition such as the role of insight in problem-solving or memory (Sternberg & Stemberg, 2012). Gestalt psychology has laid the foundations for the contemporary understanding of higher-level cognitive processes and remains robust as a descriptive framework, but like many seminal or general theories, it has not predicted or accounted for many more specific processes. Notably, the idea of cued memory retrieval through learned associations and recognition, which gisting heavily relies on, is absent and or un-emphasized in Gestalt psychology (Reyna, 2012). Perhaps this is because the Gestalt school of thought rejected the structuralist idea that complex ideas arise from association of simpler ideas. Gisting and Gestalt provide complimentary perspectives describing cognitive processes involved in making sense of complex information.

2.2.2. Narrative thinking

This process of constructing and resolving meaning is also well captured by theories of narrative thinking. Bruner distinguishes two modes of thought (Bruner, 2009). The paradigmatic mode uses rule-based processes, involving the categorization of knowledge into hierarchies and classification systems. It attempts to formulate generalizations and in this sense is context-free. It is a procedural and rigid way of thinking which breaks down when ambiguity challenges neat or orderly classification. The narrative mode, by contrast, organizes knowledge holistically by considering the emergent meaning or "story" that results from the arrangement of events, actors, temporality, purpose, and causality. Rather than formulating generalizations, the narrative mode is context-sensitive and considers particularities that are incorporated into the story being constructed by the observer. This makes narrative thinking highly flexible and able to handle contingency, anomaly, and uncertainty (Hilligoss & Moffatt-Bruce, 2014). Narrative thinking is a theoretical explanation for how humans blend memories, actions, plans, and current sensory perception into the seamless conscious experience of life. Beach argues that narrative thinking is a foundational precedent to paradigmatic thinking and is the process by which meaning is made (Beach, 2009), drawing an analogy to a novel existing merely as ink and paper until it is read by someone. The reader constructs a narrative, an imaginary world, as they read the symbols in front of them. What is currently being read is most clear, but what has been read in the past, what the reader's prior knowledge is, and what they anticipate may happen next are incorporated into a private narrative that unfolds as signs, symbols, their relationships and significance are interpreted. Just as novels feature subplots, cliffhangers, and non-chronological ordering, so too does narrative thinking involve hierarchical structures, **branching**, and **unresolved ends**. Such experiential sub-plots are micro-narratives that are evaluated and integrated into greater narratives according to how coherent they are. Often there may be multiple micro-narratives that could equally coherently integrate with the broader narrative. This defines *ambiguity*.

Analysis and interpretation of visual analytics systems involve similar processes. Visualizations present an ordered view of signs and symbols representing the world (Bertin, 1983). These are read to construct a coherent narrative and understanding. The notion of narrative is well-established in the visualization literature (Riche, 2018; Segel & Heer, 2010), however, the notion that visualization should carry any meaning beyond the data itself departs from traditionally held views of how visualizations should function. Guidelines often argue that visualization should be void of connotation, only denoting a one-to-one correspondence between data and a visual feature. In practice, however, visualizations are not purely denotative and unambiguous, they evoke associations and thus produce alternative potential meanings (Rod, 2001; Sultana et al., 2023). Just as novels unfold with subplots, loose ends, tangents, and new questions, these elements are present in the construction of personal analytical narratives when using visualizations. In work only tangentially related to this thesis, we observed how visual exploratory analyses evoke micro-narratives that tie together past experiences, knowledge, affect, and a temporal unfolding of the analysis at hand (Nowak et al., 2018). While not operationalized or further expanded upon in this research, narrative thinking provides an alternative framing to explain ambiguity in sensemaking. Whereas cognitive frames are a useful abstraction for describing sensemaking in risk-based applications, narrative thinking provides a more relatable description of sensemaking as it better matches with common every day reasoning. In addition, segmenting thinking into the narrative mode and paradigmatic mode sheds additional light on the challenges of describing complex phenomena such as avalanches in a rigorous and scientific manner.

2.2.3. Thinking with Representations

In the following section, I take a broad view of literature discussing how representations can support cognition and problem-solving activities. Here, representation goes beyond simply visual information such visual idioms and encodings. The value of visualization is often explained as reducing cognitive load and making problem solving tasks easier by leveraging human innate perceptual capabilities like “pre-attentive processing”, whereby search time to identify a target visual stimulus within a visual scene is reduced depending on the choice of visual feature used (Treisman & Gelade, 1980; Ware, 2010). While this certainly a critical component, the focus is on very low-level perceptual processes. Cognitive science discusses additional considerations as to how representations can change the cognitive work involved in problem solving at higher levels of abstraction. These considerations extend beyond the speed or efficiency with which visual information is decoded.

Researchers discuss a variety of ways that representations can support or impede thinking towards some end like a problem-solving task (Kirsh, 2010; Smith et al., 2006; Stapleton et al., 2016; D. D. Woods, 2002). The ‘representation effect’, where the representation of a problem changes the cognitive work involved in solving a problem, applies to human-machine interactions and not just static visual representations (D. Woods, 2002). The cognitive systems engineering (CSE) scholar David Woods (2002) lists several mechanisms that may influence the problem solver in ways that advance or impede problem solving:

1.) **Problem structuring**: the representation changes the nature of the problem and therefore the strategies that can be employed.

2.) **Overload/workload**: representations can shift processing to more economical forms such as simple visual processing.

3.) **Control of attention**: Representations can shift attention to what is important by taking advantage of attentional control.

4.) **Secondary tasks**: representations can introduce new tasks that shift attention.

5.) **Effort**: the effort required for a particular representation and the effort for the task at hand by proxy.

He distinguishes between representations and visual forms, highlighting how different visual forms can structure a problem in the same way and in this sense maintain the same representational form. Woods points out that researchers often “lose the forest for the trees” by fixating on low-level components of visualizations rather than the broader context of the problem space and how cognitive work should be shared between a representation and an observer towards some end. In the CSE literature this is often discussed as the difference between **coherence** (how easily a visual encoding is read) and **correspondence** (whether the relevant information and problem structure, both which go beyond simply mapping data and visual encoding, is included and supported by the interface) (Smith et al., 2006). The latter of these considers the knowledge of the observer and the problem domain, whereas the former focuses purely on data.

The advantages of representation extend also beyond the representation of data to the externalization of knowledge. Kirsh describes several dimensions for how external representations can enhance cognition (Kirsh, 2010):

1.) **Changing cost structure of inference**: The cost of sensemaking is reduced through externalization by making it more efficient and effective.

2.) **Sharing of referent**: The materialization or externalization creates a shared referent for thought between people.

3.) **Persistence of referent**: Where mental representations are seemingly more subject to alterations and harder to maintain stably, external representations are persistent and can therefore be used as a reliable structure for problem solving.

4.) **Re-representation**: Representations can be re-arranged to reveal relationships and structures that are otherwise inaccessible. A simple example is a jigsaw puzzle. This property is present in many external representations including language or mathematics. Visualization researchers have describe "free rides" as the variety or volume of inferences that can be made from a single representation (Stapleton et al., 2016). I note that this stands in contrast to common visualization guidelines that value precision (Bertini et al., 2020) and as a by-product representations that support fewer or singular types of inferences (Rod, 2001).

5.) **External representation is often closer to real structure than internal representation**: Kirsh presents music as an example. The sound of music produced by an instrument is much more a natural representation than that produced internally through reading sheet music. If the plasticity and speed of an external medium match the speed of internal thought, that external medium affords to be "thought in". This relates to why interactions support the flow of sensemaking in exploratory visual analysis (Card et al., 1999).

6.) **Affordance of computing and constructing arbitrarily complex structures**: Kirsh discusses many prominent thinkers who all have argued that the complexity of certain systems is irreducible in any coherent human (mental) way and that the best way to understand it is through modeling for approximation or using the phenomenon itself as the model.

7.) **Costs of coordinating and controlling thoughts are lowered**: External representations serve as anchors, and the actions afforded by external media are often much more effective and efficient for navigating a problem space than those by internal processes.

2.3. Ambiguity in Visual Analytics

While most visualization research has focused on uncertainty in data, many acknowledge the role interpretation plays in uncertainty more generally (Boukhelifa et al., 2017; MacEachren et al., 2005; MacEachren, 2015; Thomson et al., 2005; Zuk & Carpendale, 2007). Ambiguity is often part of the discussion, but definitions vary. For instance, MacEachren defines ambiguity as a “lack of an appropriate frame of reference” (MacEachren, 2015). Others define it in terms of the multiplicities between entities and names in data (Boukhelifa et al., 2017) or differences in interpretation between collaborators (Boukhelifa et al., 2017; Heer & Agrawala, 2008). Visualization research in natural language interfaces and mixed-initiative systems often discusses user intent disambiguation (Gao et al., 2015; Hoque et al., 2018).

Research most closely related to our own often does not use the label ‘ambiguity’. Liu et al. have a notion of cognitive alternatives describing alternative hypotheses, mental models, and interpretations in visual analysis (Liu et al., 2020). Implicit errors, errors inherent to a data set but not explicitly represented within it, are discussed in applied visualization research for infectious disease statistics (Mccurdy et al., 2019) and archaeology (Panagiotidou et al., 2021). In closely related research, Lin et al. define data hunches as “a person’s knowledge about how representative data is of a phenomenon of interest” and outline a design space for externalizing these data hunches (Lin et al., 2021). However, to date, there is little investigation into how to support and mediate ambiguity in the design of visual analytics systems.

2.3.1. The Problem of Ambiguity in Collaborative Sensemaking

Ambiguity is at the heart of challenges in shared analysis. Collaboration relies on a shared context of how evidence, artifacts, findings, hypotheses, and knowledge relate (Soares et al., 2016) to establish common ground (Yusoff & Salim, 2015), a concept adapted from linguistic and social psychology describing shared understanding enabling communication (Heer & Agrawala, 2008). However, as communicating analysis demands effort beyond the task of analysis itself, collaborators follow the principle of least collaborative effort exerting the least possible effort for communicating a message (Heer & Agrawala, 2008). This limits how much information is captured and as soon as analysis is shared, some context is lost. Communications will be ambiguous particularly when using spoken or written language, a symbolic system of representation which naturally invites multiple interpretations (Sennet, 2016). It may not be apparent how evidence relates to analytic findings, how prior knowledge was used, what work has been done as well as what remains.

These challenges are especially pronounced in asynchronous sequential hand-off of analysis. Many domains involve shift-changes where analysis is shared across successive work shifts involving different collaborating analysis. As knowledge work like analysis involves emergent findings that can be difficult to articulate, hand-off is disruptive and can be a considerable challenge In many domains, but especially those with high consequences for failure (Patterson, 2008; Patterson et al., 2004, 2007, 2017; Patterson & Woods, 2001). Sharing analysis is often a final step when individuals are fatigued, which can further degrade the quality of communication (Sharma, 2008). In addition, it may be difficult to anticipate which information will be relevant to collaborators in the future (Patterson, 2008). Partial findings or formative sensemaking processes that may be relevant or critical to share are often omitted as they are difficult articulate, and without the context within which sensemaking was instantiated, the nature of the problem will appear ambiguous (Sharma, 2008).

Researchers have suggested a set of solutions to ease the burdens of collaboration. Generally, these involve the capture of critical information during analysis rather than after, and ways to structure captured information for communication. Many have suggested the use of visual annotations (Andrienko et al., 2018; Patterson & Woods, 2001; Zhao et al., 2018) , which serve to point to important information in context and externalize knowledge about its relevance to analysis. However, on their own, knowledge externalizations such as annotations lack structure and may become visually cluttered and difficult to make sense of. This presents design challenges for how to organize the meta-data embedded through annotations or markup to make it tractably retrievable and navigable.

This is why schematization is often highlighted as a key support mechanism for collaboration (Chen et al., 2011; A. P. Fischer, 2011; Mahyar & Tory, 2014; Mathisen et al., 2019; Zhao et al., 2018). Domain knowledge may be used to structure and scaffold externalizations, providing direction to analysis, what is captured, and how easily it can be understood by collaborators (Andrienko et al., 2018). However, rigid standardized protocols for hand-off often fail to capture critical information that is difficult to articulate (Hilligoss & Moffatt-Bruce, 2014). This is why such schematization mechanisms should be flexible and editable to serve the needs of the situation (Andrienko et al., 2018; Federico et al., 2017; Wright et al., 2006). Researchers suggest the use of flexible templates, to represent a structural model of a problem space (Wright et al., 2006). By filling a template with thoughts, observations, notes and links to evidence, such unstructured media are schematized easing the burden of gathering and distilling such information.

However, as a lot of expert analysis involves tacit knowledge (knowledge employed without being conscious of it), schematization demands additional introspection to make the tacit explicit (Shipman & Marshall, 1999). In doing so, it necessarily disrupts the task at hand. In such circumstances or when faced with situations of high uncertainty, ambiguity may actually serve a productive purpose as it does not demand additional effort in articulation and it can accurately convey the state of current understanding (Sterner, 2022). Such ambiguous information within the context of shared work environment can help collaborators cue-in on information in ways that rigid hand-off protocols do not (Mueller et al., 2006). Such information can help collaborators become aware of each other's activities and coordinate work (Heer & Agrawala, 2008). Whereas collaborators in a shared physical environment can see movement, gestures or how physical objects have been manipulated, virtual collaborative environments must represent such explicitly to support awareness (Marriott et al., 2018). Various mechanisms such as color coding the work of collaborators according to their identity (Drouhard et al., 2017; Wu et al., 2013; Zhao et al., 2018) or reconstruction and exploration of past analysis activities (Malik et al., 2011; Wu et al., 2013; Zhao et al., 2018) have been employed towards this end.

Existing Work in Collaborative Visual Analytics

Supporting collaborative data analysis is a key research challenge in visual analytics (Heer & Agrawala, 2008; P. Isenberg et al., 2011; Mahyar & Tory, 2014). Researchers in Collaborative Visual Analytics (CVA) have long recognized that visualizations have the advantage of making “deictic” references relevant materials providing shared context and disambiguating meaning (Heer & Agrawala, 2008). In the CVA literature, these types of knowledge visualizations have traditionally taken the form of concept maps with hypermedia as deictic pointers to reference material such as evidence (Chung et al., 2010; Goyal et al., 2014; Mahyar & Tory, 2014; Shrinivasan & Van Wijk, 2009; Zhao et al., 2018). These have been used to support both synchronous and asynchronous collaboration, coordinating work by highlighting the work that has already been done as well as providing references to relevant materials.

Such knowledge visualizations are also helpful in collaboration by clarifying where understanding is in conflict – highlighting areas where the representations does not match the viewers mental model. This is a particularly useful property as it can direct attention to discrepancies and elicit knowledge that is relevant to resolve the discrepancy (Hoffman et al., 2006). This use of concept maps has proven useful in negotiation (Swaab et al., 2002), problem solving (F. Fischer et al., 2002), and eliciting explicit mental models of weather forecasting procedures (Hoffman et al., 2006).

The CVA literature provides several examples of strategies employed to make the capture, communication, and retrieval of knowledge findings less burdensome and more effective. Strategies employed have ranged from automatically importing annotations from individual workspaces into a shared workspace (Goyal et al., 2014), representations of planned courses of action in relation to the analysis environment (Prue et al., 2014), representing interaction logs of navigation paths alongside data visualizations and knowledge produced during analysis (Shrinivasan & Van Wijk, 2009), hierarchically-structured annotations describing analytic questions linked to application states (Mathisen et al., 2019), and combining automated analysis capture through interaction logging combined tags categorizing and structurally relating different tacit components of analysis such as unresolved, uncertainties, perceived relationships and hypotheses (Zhao et al., 2018). Such approaches, fusing the analytic process with reporting of analysis, reflect the logic behind literate approaches to computing (Mathisen et al., 2019; Rädle et al., 2017; Wood et al., 2019), which leverages the understanding that narrative structure is inherent part of cognitive work such analysis which can naturally be repurposed for communication. It should be noted that many of these tool's focus on exploratory visual analysis, often grounded in intelligence analysis, with emergent findings. There is comparatively less research of how CVA tools can be tailored to address needs of monitoring in critical and resource constrained applications (Malik et al., 2011; Wu et al., 2013).

Moreover, visual analytics systems have been heavily focused on supporting bottom-up data-driven analyses and have neglected the broader picture that includes explicit support for knowledge-based top-down processes (Andrienko et al., 2018; Choi, Childers, et al., 2019). Researchers have argued that visual analytics should consider the reality of the domain of application as well as knowledge constructs of the domain of application to guide the design of visual analytics systems (Andrienko et al., 2018). Furthermore, scholars have argued that the goal of visual analysis is not the generation of insights, but rather the process of iterative development and calibration of mental models to some end, including assessment, prediction, or a decision (Andrienko et al., 2018).

2.4. Summary

There has been little investigation into how to support and mediate ambiguity in the design of visual analytics systems (Lin et al., 2021; Liu et al., 2020; Mccurdy et al., 2019; Panagiotidou et al., 2021). There is a need to better support knowledge-based processes in such systems (Andrienko et al., 2018; Choi, Childers, et al., 2019) both as they pertain to individual and shared sensemaking around data. This is especially important in applications involving risk because sensemaking then involves constraints, conflicting or ill-defined goals, distributed work, and decision-making that extend beyond the analysis of data in isolation (Hollnagel & Woods, 2005; Johansen & Rausand, 2015; Smith & Hoffman, 2017). To the best of my knowledge, no visual analytics research has specifically focused on ways to support and mediate ambiguity in sensemaking within complex risk prediction and management applications.

Chapter 3. Problem Domain

In the following section, I provide relevant background information about the problem domain of avalanche forecasting. I discuss the formal definition and purpose of avalanche forecasting and its role in broader risk management practices, review existing literature describing how forecasters reason and use evidence, describe the specific and distinct context of public avalanche forecasting, and discuss existing literature related to ambiguity in this domain. My research is applied within Canada and consequently I describe the practices and theoretical concepts from within this industrial context. While different practices exist internationally, the differences are minute and the general concepts, practices, and challenges faced are similar. I provide additional background information about Avalanche Canada, contextual information about the organization and our engagement, and a discussion of the sample of participants involved throughout this research.

3.1. Overview, Definition and Purpose

Avalanche forecasting is a **collaborative hazard and risk analysis, prediction and communication** activity relying on **continuous monitoring and review** distributed across individuals and organizations. Forecasters are responsible for communicating their assessments of avalanche hazards as part of broader risk management procedures. Forecasting occurs in a variety of operational avalanche safety contexts such as ski resorts, commercial helicopter skiing operations, operations overseeing avalanche safety in transportation corridors, or remote office-based public avalanche forecasting among others. Depending on the context, these assessments may be communicated to the public, professional avalanche safety operators, or used in personal risk management.

Snow avalanches are natural disaster phenomena where structural weaknesses in the layered seasonal snowpack in layers of snow are triggered releasing destructive masses of snow that endanger human life and infrastructure (McClung, 2002a). Triggers may be environmental factors, such as the weight of new snow, or human triggers, such as the weight of a skier or snowmobile. Avalanche forecasting is formally defined as “the prediction of current and future snow instability in space and time relative to a given triggering level" (McClung, 2002a).

Avalanches are a highly dynamic and complex natural phenomenon. Weather systems interact with mountainous terrain producing variation in atmospheric conditions often at a very localized spatial scale (Ahrens & Henson, 2018). A multitude of physical processes affect the formation, evolution, and stability of snowpack over the course of a season (McClung & Schaerer, 2006). As avalanches can be triggered by and impact humans and human infrastructure, human behaviour is also an important element that interacts with avalanche conditions and further contributes the complexity of avalanches.

This complexity results in considerable uncertainty such that the goal of avalanche forecasting is to "minimize uncertainty about instability" (McClung, 2002a), where uncertainty is defined as the "state (even partial) of deficiency of information related to understanding or knowledge of an event, its consequences or likelihood" (Campbell et al., 2016).

3.1.1. The Role of Avalanche Forecasting in Avalanche Risk Management

The Canadian Avalanche Association (henceforth CAA) uses the ISO 31000 Risk Management Principles and Guidelines (International Organization for Standardization, 2009), a general risk management guideline setting standards across any risk management application, as a framework for risk management (Campbell et al., 2016). Avalanche risk management is composed of three stages including establishing the context which determines the scope (objectives, hazard/risk criteria, and relevant factors of activities undertaken) and situation (elements at risk such as people or infrastructure, potential risk scenarios, and the spatial and temporal scales), hazard and risk assessment which involves the identification, analysis and evaluation of the present hazard and risks, and risk treatment which involves risk control and mitigation strategies. Depending on the context and the risk management needs, this is either implemented using long-term planning (e.g., construction of critical infrastructure) or employing short-term operational avalanche management activities (e.g., commercial backcountry mountain guiding). Avalanche forecasting falls into the risk assessment stage of operations, but as the process of avalanche risk management is an iterative process involving continuous monitoring and review, all components of the avalanche risk management process may bear relevance to avalanche forecasting.

Risk assessment begins with hazard assessment which treats the potential of any hazard (such as an avalanche) as independent from the risk it poses to people, organizations, or infrastructure, technically referred to as an “element at risk”. Hazard assessment typically concludes with an avalanche hazard forecast that characterizes avalanche hazards according to various dimensions which I describe below. This is then applied to a given element-at-risk by considering its exposure and vulnerability to the identified avalanche hazard. Risk management decisions (e.g., deliberately triggering avalanches in a controlled manner using explosives) are based on that assessment. Avalanche forecasting may often just involve hazard assessment but is often tied together with broader risk assessment processes in different operational settings. The operational context determines the types of observations available for avalanche forecasting, the ways in which such information is processed, the tools available, and how such information and activities fit in the broader risk management process. As will be expanded in a later section, public avalanche forecasting is concerned with communicating hazard assessments to the public and does not formally consider elements at risk and risk management.

3.2. Hazard Assessment

The Conceptual Model of Avalanche Hazard (henceforth CMAH) formally defines the essential components of avalanche hazard and streamlines them into a procedural guideline and workflow for hazard assessments independent of the element at risk (for example people or infrastructure) (Statham, Haegeli, et al., 2018). This separation ensures a common standard for assessments despite the differences in the context within which forecasting is applied. For instance, managing avalanche risks will be different for guests at a ski resort than for traffic along a mountain highway corridor. The CMAH has become a formalized and explicit part of the forecasting workflow in information and technology systems around North America (Haegeli et al., 2014; Statham et al., 2012), and is an important component of how avalanche hazards are communicated to the public. While the CMAH is primarily used in North America, it is a useful model to describe how observations are analyzed to arrive at hazard assessments in any country/setting.

The CMAH is built upon the concept of an **avalanche problem** and **avalanche problem types,** which describe identifiable, recurring patterns in avalanche hazard with respect to the meteorological conditions that lead to their formation, the nature of the commonly associated avalanche activity, the typical evolution of conditions, and effective mitigation strategies (Statham, Haegeli, et al., 2018). Avalanche problems are instances of avalanche problem types presenting operational threats which avalanche forecasters describe with respect to their avalanche character/type, location, likelihood, and size. The model prescribes a sequence of questions for analyzing this information: (1) What type of avalanche problem(s) exist? (2) Where are these problems located? (3) How likely is it that an avalanche will occur? (4) How big will the avalanche be? The avalanche problem type sets expectations about what information may be relevant and directs the information-seeking process. As terrain influences how snowpack and resulting instabilities form, forecasters then identify the terrain locations (described using common terminology such as elevation, aspect, vegetative cover, or geomorphology) where this hazard may be found. Next the forecaster determines how likely an avalanche is to take place by considering both how sensitive an instability is to triggering (derived from observations such as past avalanches and field tests) and it's spatial distribution (considering how easily evidence is found within the set of identified terrain locations). The forecaster then considers the potential destructive size of the avalanche. The questions in the CMAH are answered using qualitative ordinal categories that are at the discretion of the subjective judgement and interpretation of the avalanche forecaster to determine. There is a formal and prescriptive method of how ordinal categories are combined to higher levels (such as how spatial distribution and sensitivity combine to determine likelihood). The output of this process describes the avalanche hazard by combining the ordinal values of likelihood and size into a matrix visualization (Fig. 1). Uncertainties are expressed and shown as ranges spanning the possible values of each ordinal dimension. When plotted, this results in a rectangular shape that provides a common visual representation for characterizing an avalanche hazard and its identified uncertainties.

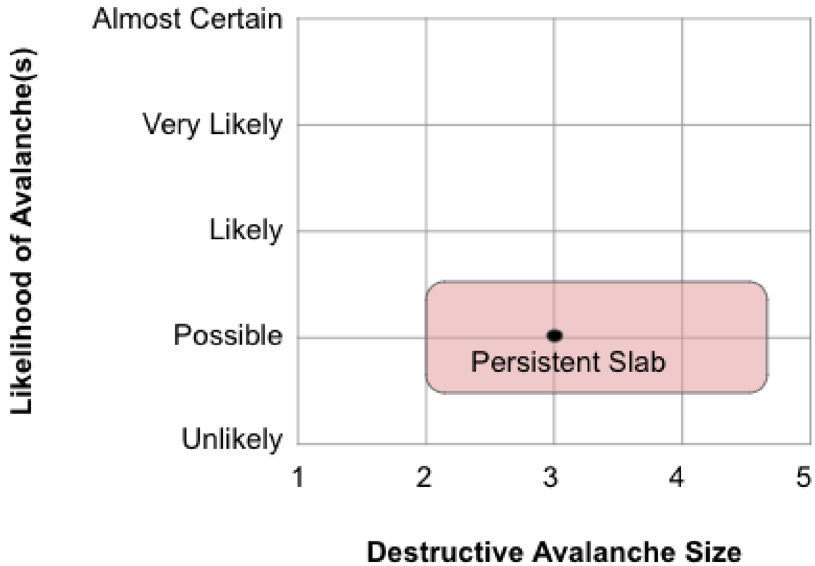


Figure 1 **Matrix visualization showing the size and likelihood ranges of a given avalanche problem.**

These formalisms provide a common language for forecasters to express their mental models and understanding of avalanche conditions more consistently. However, inconsistency remains a challenge as the uncertainty of forecasting and the nature of evidence demand the use of personal knowledge and experience to subjectively judge hazards along these scales and dimensions. I expand on this in the following sections.

3.2.1. Uncertainty in Avalanche Forecasting

The “goal of avalanche forecasting is to minimize uncertainty about instability” (McClung, 2002a). In the avalanche context, uncertainty is the “state (even partial) of deficiency of information related to understanding or knowledge of an event, its consequences or likelihood” (Campbell et al., 2016). The avalanche safety literature commonly distinguishes between two primary types of uncertainty: Aleatoric uncertainties refer to the inherent random and natural variability of complex systems which cannot be reduced (rather it can only be described probabilistically), and epistemic uncertainties refer to a lack of knowledge that could be known in practice but is not (Campbell et al., 2016; Jamieson et al., 2015). This distinction is relevant as it describes how forecasters incorporate probabilistic assessment to address aleatoric uncertainties or seek further information to reduce epistemic uncertainties.

McClung describes three primary sources of uncertainty in avalanche forecasting: (1) The spatiotemporal variability of snow cover and the influences of terrain; (2) Incremental changes from snow or weather conditions; (3) Human factors stemming from variations in human perception and estimation (McClung, 2002a). As human forecasters conduct the assessment process, their perception of reality underlies all these sources of uncertainty. McClung explains that the only way to reduce uncertainty in avalanche forecasting is to seek new information of the right kind or to take actions that deal with variations or resolution in human reasoning. What is meant by information of the right kind is information that can reveal something new about instabilities in the snow. Due to the complexity of the phenomenon and a dearth of data for statistical-style analyses (random sampling) in any realistic operational setting, information about stability holds considerably less diagnostic value than novel information about instabilities. However, as there are identifiable patterns in conditions that create a particular avalanche hazard, one can form expectations about where and how to seek information (targeted sampling). Moreover, the cost of a false negative (predicting no avalanche when or where one does occur) is much higher than the cost of a false positive (predicting an avalanche when none occurs) further steering investigation and sampling of evidence in a targeted direction towards signs of instability. This describes how observations and data are gathered for avalanche hazard assessment.

3.2.2. Observation Types, Information Sources and Systems

Avalanche forecasting relies on a variety of information sources and data streams. As this research addresses public avalanche forecasting in Canada, I will describe the types of data and systems forecasters have available in this context. These sources are not exclusive to public forecasters but are the primary data and systems of concern for this research. Again, while the specifics of how such data are defined and disseminated may vary internationally, the general approach is very similar.

Forecasters rely heavily on point observations sourced at specific spatial locations by humans or sensors such as those on weather stations. These observations are used to understand current and historic avalanche conditions and form a significant portion of data used in hazard analysis. These point observations are integrated with the prior day's forecast and assessments (LaChapelle, 1980) using expert subjective judgment and knowledge to create a ‘nowcast’, an assessment of current instabilities and avalanche conditions (McClung, 2002a). This is then integrated with numerical weather forecast models to predict how future atmospheric conditions will affect the snowpack and in turn avalanche conditions (McClung & Schaerer, 2006).

In the following sections, I describe the data streams available for avalanche forecasting in the Canadian context. I describe the type of data available, how they are gathered, the system it is presented within, and any specific challenges this presents. These include observations made in the field by professionals or recreationists, observations about meteorological conditions from weather stations, and various products for viewing numerical weather prediction models.

Industry Information Exchange (InfoEx)

The Industry Information Exchange (InfoEx) is a web platform and database where Canadian professional avalanche safety operations share daily reports relevant to avalanche safety with each other (Haegeli et al., 2014). Subscribers to this system include ski resorts, backcountry guiding operations, highway and railway avalanche safety operations, avalanche safety consultancies, individual mountain guides among many others. These subscribers report “observations” gathered from the field areas within which they work as well as formal hazard assessments. Hazard assessment include various components of avalanche problems as described in the CMAH and higher-level avalanche danger ratings (B. Lazar et al., 2016). Reported observations data are entered using web entry forms organized according to different types of observations and presented corresponding data tables[1]. Observations are reported following the Observation Guidelines and Recording Standards (OGRS) set by the Canadian Avalanche Association (1995). These include technical information about weather such as “Field Summaries” providing a general summary overview of weather conditions in the subscribers operating tenure or “Weather Observations” which provide specific weather conditions at a permanent weather site at regular time intervals. “Snow Profiles” provide technical information about snowpack structure and stability gathered by professionals by digging ‘pits’ in the snow. Finally, “Avalanche Observations” provide records of observed specific avalanche occurrences and details about their

[1] http://infoexhelp.avalancheassociation.ca/wiki/Documentation_overview

characteristics, while “Avalanche Summaries” provide free-text summaries of avalanche activity across an entire area that an operator made observations in for a particular day. As “Avalanche Observations” and “Avalanche Summaries” are an essential part of this research, I will expand on these below.

Individual “Avalanche Observations” records or data entities correspond to observed avalanche of the same type. The is “character” (avalanche type) field to specify this. There are a number of other fields describing the dimensions and characteristics of the observed avalanche(s) such as the destructive size of the avalanche or what triggered the avalanche(s). Many of these data are presented as ranges of numerical values based on the observing operators estimates. When more than one avalanche is being reported, the number of avalanches observed is specified using either a numerical value (e.g., 1-50) or a qualitative ordinal value (e.g., “several”, “numerous”). Photographs of the observed avalanches are included. In addition, there is associated location data which may include one or more spatial vector polygons or point locations corresponding to where avalanche observations were made. Note that these present varying levels of spatial resolution for reporting one or more avalanche observations. A final noteworthy field is the observation date. As avalanches may be observed in motion or as avalanche debris after the fact, avalanche observations often include estimates of how old the debris is which can then be used to infer an estimated time at which the avalanche occurred.

InfoEx subscribers vary widely in operating procedures and needs. In addition, they are under considerable time pressure from existing work leaving less time for data entry. The InfoEx system was design to be adaptable and flexible to address these issues (Haegeli et al., 2014). Consequently, the way data are reported and how they may be interpreted will depend on the specific operational context from which they are sourced. Again, as these observations are gathered using a targeted sampling approach, interpreting how observations come to represent conditions more broadly requires integrating an understanding numerous contextual factors contained in multiple attributes such as terrain characteristics or locations where avalanches were observed.

The InfoEx system also allows subscribers to view and analyze through a variety of views including data tables, maps, and visualizations (Fig. 2).

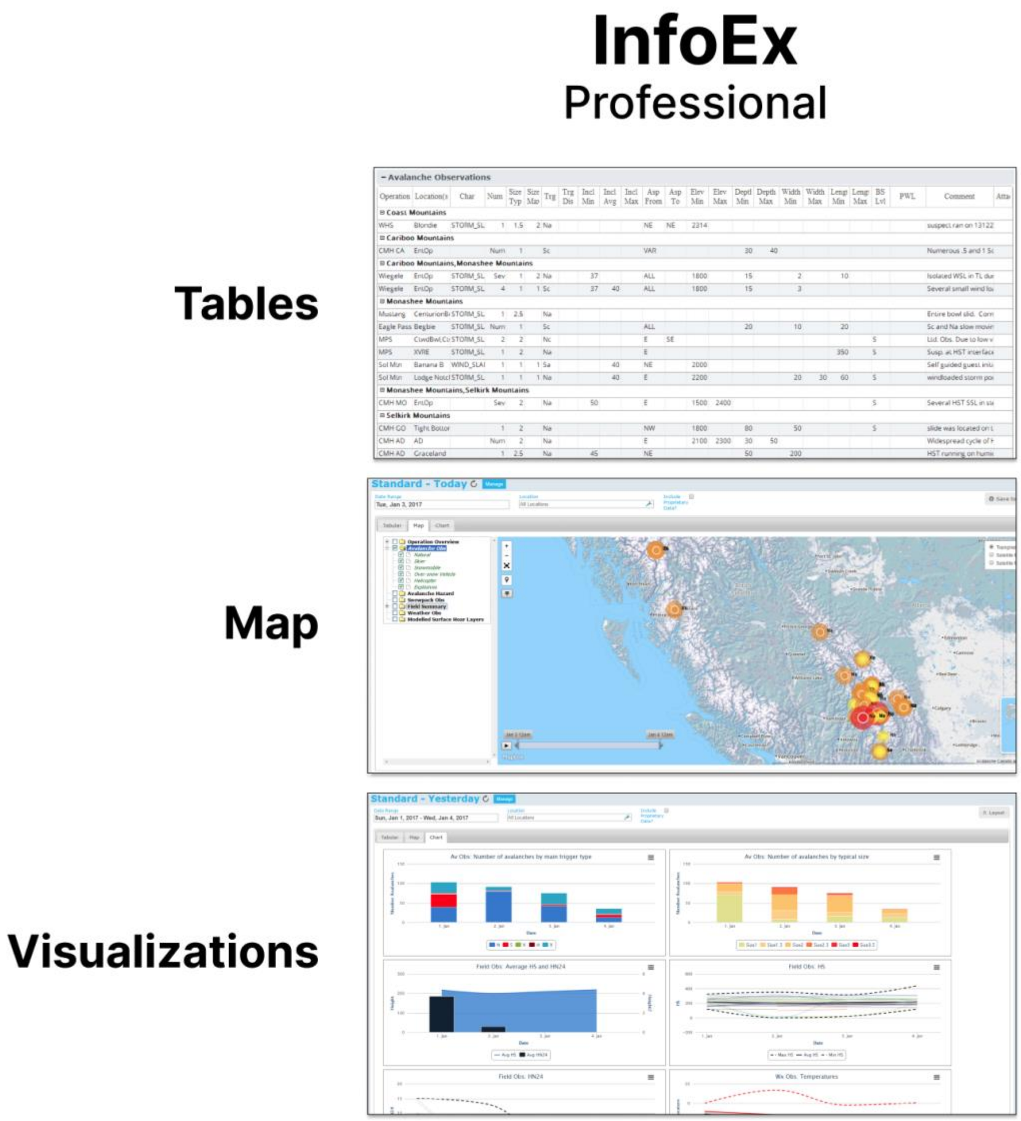


Figure 2 The data table, map, and visualization views available in the InfoEx system used for sharing reports amongs avalanche professionals.

Mountain Information Network (MIN)

Recreationists also share trip reports using the Mountain Information Network[2] provided by Avalanche Canada, a public avalanche forecasting agency. Similar to the InfoEx, a variety of information corresponding to observations made in the field are submitted through structured web forms. These include information about weather

---

[2] https://avalanche.ca/mountain-information-network

conditions, snowpack structure, observed avalanches, incidents involving avalanches, or "quick" summary information about conditions. As with the InfoEx, these data are composed of structured fields, free-text fields, and images.

These reports are publicly available on Avalanche Canada's website and can be accessed alongside other information such as the daily avalanche bulletin (Fig. 3). The two ways these reports are navigated are through a map showing the spatial locations of each observation report and a list of reports. Clicking a report on the map opens a sidebar with report details that can be further expanded to a fullscreen view. Similarly, a fullscreen view of report details is accessed by navigating from the list view.

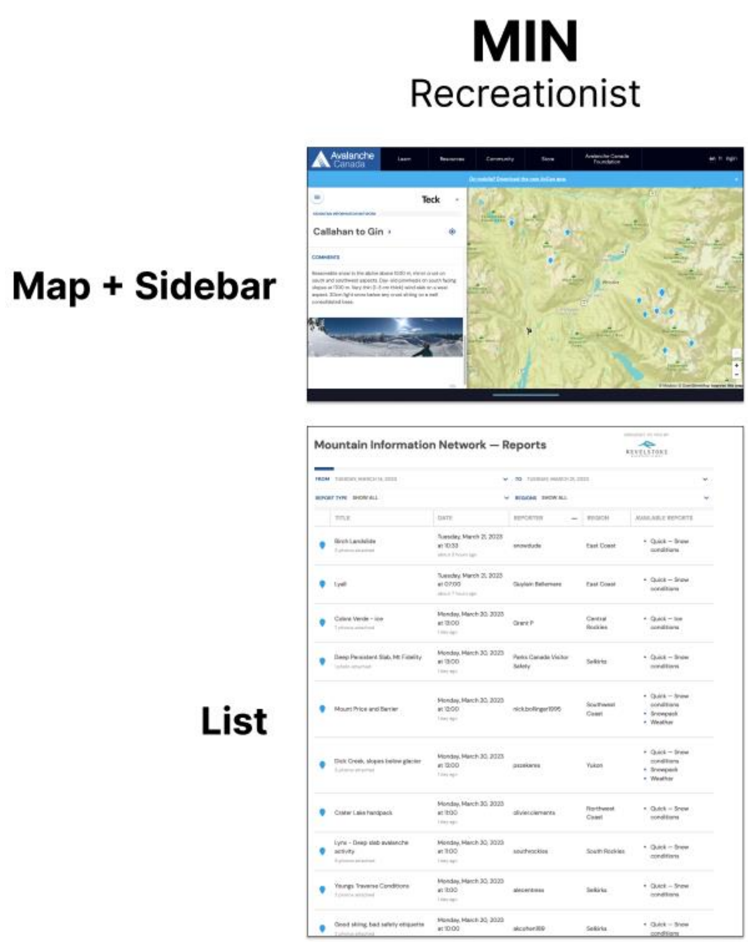


Figure 3 The primary map and list view used to access crowdsourced reports from recreationists on Avalanche Canada's Mountain Information Network (MIN) system.

Weather Stations

Weather stations provide a critical data source for understanding mountain weather. They are placed at various elevations and in locations that are representative of broader spatial areas. They provide sensory telemetry readings for snow depth, the mass of snow measured as a snow water equivalent, liquid precipitation, wind speeds and directions, and temperature among others. Forecasters use weather stations as a source of “ground-truth” to compare weather model predictions against. In Canada, there are a variety of weather station network providers accessed through a variety of websites and portals. ARFI, the Avalanche Research Forecasting Interface[3], gathered links to access individual weather station data as well as other forecasting resources in a map-based interface (Fig. 4). Clicking weather station icons navigates to other webpages where real-time, near-real time, and historical telemetry readings may be found in a variety of formats. Primarily, these are available in text tables, but charts for individual stations or sometimes available as well.

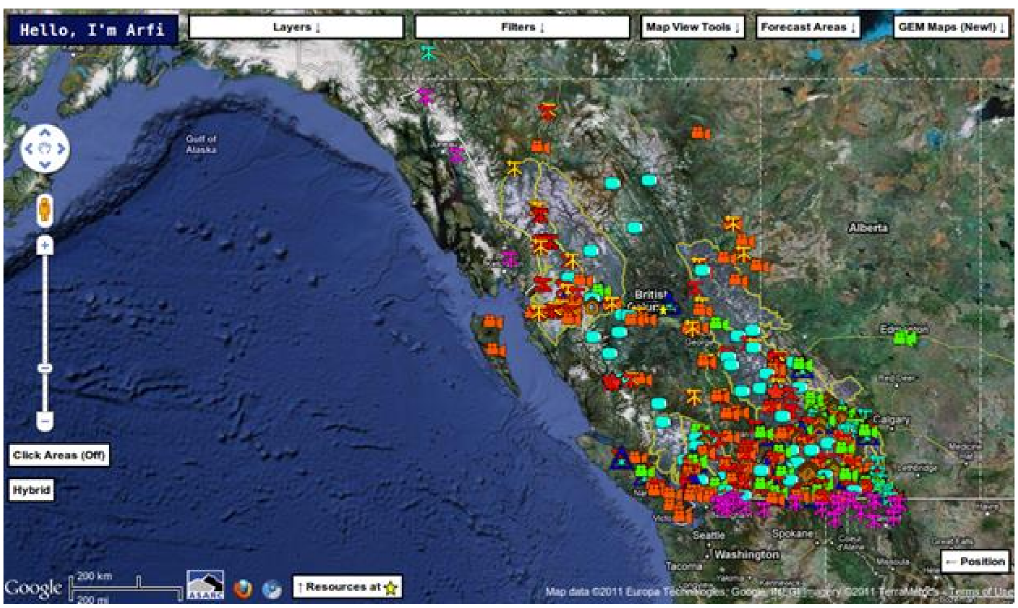


Figure 4 **The Avalanche Research Forecasting Interface (ARFI) used to access weather station telemetry data among other relevant resources.**

---

[3] ARFI.avalanche.ca

Weather stations are sparsely distributed relative to the spatial areas they are used to represent (Lundquist et al., 2019) leaving forecasters to use personal knowledge of how local terrain and weather systems interact to extrapolate from these point data. In addition, weather stations are subject to a variety of sensor and transmission errors caused by environmental factors. Presently, there is no comprehensive automated quality assurance procedure that accounts for all possible errors in these data (Mekis et al., 2018). Diagnosing errors and how individual weather stations come to represent broader weather patterns is a matter handled through the forecaster's judgment and interpretation.

Numerical Weather Prediction Models

Meteorological weather models form a significant part of the information avalanche forecasters investigate. After assessing current avalanche conditions using observations made from a variety of sources such as those discussed above, predictions about future conditions can be made by considering weather forecasts. Technical considerations and the variety of products used in avalanche forecasting are beyond the scope of this research. However, it should be noted that mountain weather is highly complex and that the Meteorological Service of Canada (MSC) provides a dedicated mountain weather product which is disseminated through Avalanche Canada's website[4].

3.2.3. Reasoning About and Weighing Evidence

To describe how avalanche forecasters reason using evidence, it is useful to differentiate between inductive, deductive, and abductive reasoning (Douven, 2017). Deduction involves deriving a conclusion about a set of particulars that is *necessarily* true based on general or universal premises. For instance, if it takes 15 minutes to get to 11:00 AM appointment, one can deduce that one needs to leave prior to 10:45 AM. Induction, by contrast, involves generalizing based on particulars in a way that involves probability or uncertainty. For instance, one may expect that the most popular ice cream flavour at an ice cream shop probably tastes good and would likely be a good purchase. Abduction is a specific form of inductive reasoning where one makes an inference to the best possible explanation from among a set of potential alternatives. For instance, if one

[4] Avalanche.ca/weather/forecast

observes scratching and squeaking noises coming walls and droppings on the floor, one can infer that their house probably has a mouse or rat infestation.

The continuous act of seeking new and relevant information to address uncertainties characterizes the predominant inductive reasoning processes in avalanche forecasting (LaChapelle, 1980; McClung, 2002a). While not widely discussed in the literature, inductive reasoning is most often of the abductive variety because forecasters explain observations using a set of common constructs and probabilistic considerations while making inferences to the best possible explanation. Due to the complex and dynamic nature of avalanche phenomena, new information must continuously be sought to update understanding of the state of potential hazards and risks. This probabilistic style of thinking is frequently compared to Bayes Theorem. Deductive reasoning is also involved when considering physical deterministic processes, computational modeling, or decision aids.

LaChapelle and McClung describe the process of how different types of observations are ranked or weighed according to “informational entropy”, a concept adapted from information theory (LaChapelle, 1980; McClung, 2002a). Low entropy Information is the most valued as it is subject to lowest uncertainty, is most easily interpreted by humans, and is most closely related to the phenomenon in question. In spite of this, relatively higher entropy information may outweigh lower entropy information if it reveals new information about instability (McClung, 2002b). This is related to the concepts of strength (inverse of entropy) and weight of information described by Vick (Vick, 2002). Vick cautions against overconfidence from overuse of strong information sources with little weight or predictive value, and conversely, warns against underconfidence from lack of consideration of less strong but more weighty evidence. Based on this logic, observations used in avalanche forecasting are grouped into three classes ranked by their levels of informational entropy (McClung, 2002b). Class III data (high entropy) are meteorological data such as weather forecasts or reports about atmospheric conditions. Class II data (intermediate entropy) are related to snow structure such as stratigraphy, snow temperature, and characteristics of snow crystals. Class I data (low entropy) are data that relate directly to the mechanical properties of or evidence about instabilities in the snow. Evidence that may belong to a higher entropy class may outweigh evidence in a low entropy class if it reveals positive information about an instability.

3.2.4. Public avalanche forecasting

The objective of a public avalanche forecasting operation is to produce a daily public avalanche bulletin for a particular area communicating hazard assessments (McClung, 2002a). This bulletin serves as the starting point from which recreationists base their personal risk management procedures when traveling in the backcountry. This means that public avalanche forecasters conduct hazard assessments independent of any risk analysis which would consider a specific element at risk.

Public avalanche bulletins generally present information in a tiered structure ordered in increasing levels of complexity and decreasing level of abstraction[5]. The starting point is the ordinal five-level danger scale (Statham et al., 2010). Danger ratings are assigned to different elevation bands for several days in the future. Next, the day's avalanche problems and their characteristics, as described in the CMAH, are presented. A combination of Icons and text is used to present danger ratings and elements of avalanche problems. Following these, summary text of weather conditions, snowpack structure, and recent avalanche observations is presented. This structure ensures that information is accessible, concise, yet allows investigation of more detailed information depending on the level of expertise a reader might have.

Evidence in Public Avalanche Forecasting

The operating context of avalanche forecasting determines the types of observations available and how they are used. Three relevant spatial scales are adapted from meteorology for avalanche forecasting: synoptic (forecasting for region or mountain range); meso (forecasting for highway avalanche area or ski area); micro (forecasting for an avalanche path or specific terrain feature) (LaChapelle, 1980; McClung, 2002a). Public avalanche forecasters operate at meso to synoptic spatial scales with the objective of producing daily public avalanche bulletins often providing information about avalanche hazards several days into the future (McClung, 2002b). This low resolution spatial and temporal scale results in significant variability and leaves public avalanche forecasters to rely relatively more heavily on meteorological class III data than other

[5] https://www.avalanches.org/standards/information-pyramid/

avalanche safety operating contexts. However, class II and class I data are a critical part of avalanche hazard assessment in public avalanche forecasting.

3.3. Ambiguity in Public Avalanche Forecasting

The avalanche safety literature discusses several issues that are related to ambiguity and the complexity of the domain.

As all avalanche forecasting relies on the expert judgment and knowledge of individuals, it is subject to variations in perception (McClung, 2002a). In the context of public avalanche forecasting, these variations have become apparent through research investigating varying consistency in public bulletins (Clark, 2019; Hordowick, 2022; B. Lazar et al., 2016; Statham, Holeczi, et al., 2018; Techel et al., 2018). Avalanche problem types are not mutually exclusive or independent and how observations are used to determine specific avalanche hazard assessments is not well defined. Consequently, choices involved in which to report or how to transition between them is a source of uncertainty and confusion (Hordowick, 2022; Klassen, 2013; Klassen et al., 2013). Among other factors such as physical differences between different regions forecasters operate in, inconsistencies are attributed to cultural and procedural differences in forecasting organizations, unique perspectives/biases of any individual forecaster, and varying perceptions on risk communication priorities, and varying approaches of how deal with uncertainty. Unique and varying perspectives at the levels of individuals, organizations, and target audience are a source of ambiguity and manifest into issues of inconsistent hazard assessments in public bulletins.

Ambiguity is a pervasive aspect of forecasters work constituted by the complexity of the phenomenon, the nature of available data, and the challenges of characterizing and communicating this complexity. These issues of ambiguity are common in risk assessment (Johansen & Rausand, 2015) and applications such as the prediction of other natural disasters and extreme weather events (Beven et al., 2018).

According to Maguire & Percival (2018), existing tools and procedures in the avalanche domain do not capture all the cognitive work done by forecasters. This suggests that visual analytics solutions could be of benefit and demonstrates how this domain is an ideal application in which to investigate visual analytics solutions targeting knowledge-based processes and ambiguity in sensemaking. In addition, the forecaster's heavy reliance on prior knowledge to interpret data sampled in a targeted manner has many parallels to the challenges described in visualization research on alternatives (Liu et al., 2020), data hunches (Lin et al., 2021), and implicit errors (Mccurdy et al., 2019; Panagiotidou et al., 2021). The similarities of these challenges and distinctions in the characteristics of the application domain suggest that visual analytics research on ambiguity in public avalanche forecasting may extend this literature by offering novel insights offering novel insights and alternative solutions.

3.4. Avalanche Canada

This research is conducted in collaboration with Avalanche Canada[6], a non-government and non-profit public forecasting organization. This research began as part of a broader engagement to help develop Avalanche Canada's visualization and analytics infrastructure. Avalanche Canada issues daily public bulletins throughout the winter season for most of western Canada. This organization is also involved in a variety of avalanche research projects and coordinates, develops, and delivers avalanche awareness and education programs. As this study ran over the period between 2019-2023, the exact size of the organization has fluctuated. However, there are roughly 12-18 forecasters, 12-18 operations staff, and 12-18 field technicians. These avalanche professionals are from various backgrounds and varying levels of expertise. I note that the organization and avalanche industry is heavily skewed toward self-identifying males. The participant sample gathered in this study, with 4 out of 18 participants self-identifying as female, is therefore representative of the organization and broader domain. I provide further relevant details about participants in each study and provide a unified participant table below for reference (Table 1). Meta-data for each participant are presented if they were gathered within the respective studies, they participated in.

[6] Avalanche.ca/about

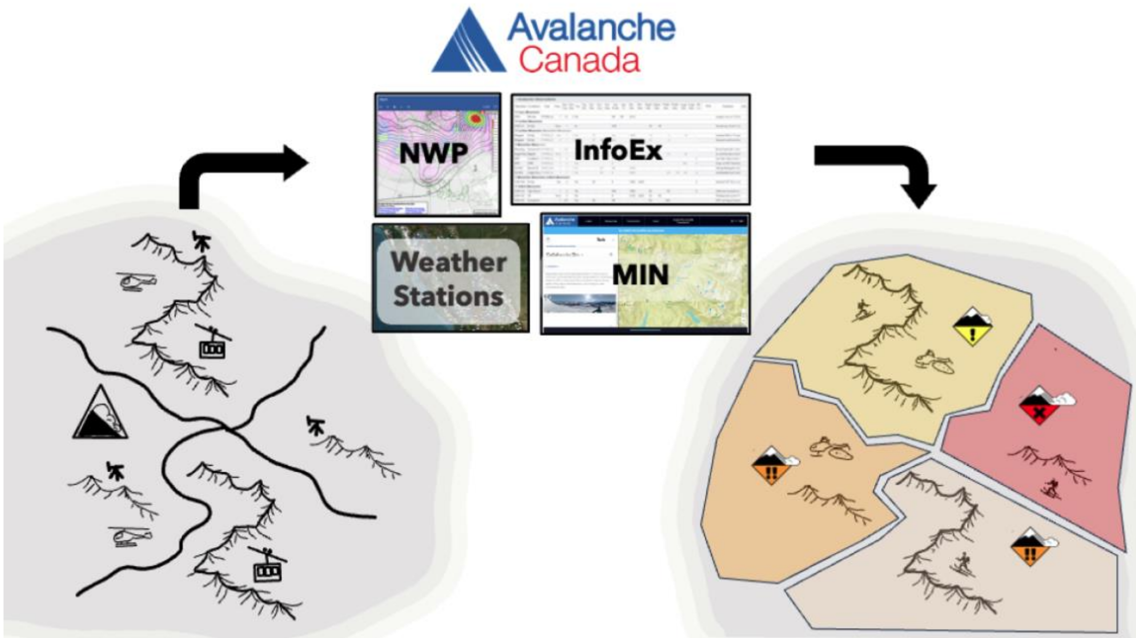


Figure 5 **A process diagram of information sources, tools and products at Avalanche Canada. Forecasters draw on point observations from remote weather stations and human field-reports. These are analyzed using disparate tools to summarize regional conditions.**

As compared to other public forecasting organizations in Canada such as Parks Canada[7], a government agency overseeing various functions in Canadian national parks including avalanche safety, Avalanche Canada's forecasters work in a remote, office-based setting. The land area Avalanche Canada is responsible for forecasting is much larger than any other in the world, and while field teams have increasingly been deployed in areas where reports from professionals and recreationists are scant, this remote setting dictates much of this organization's approach (Fig. 5). This influences the available data and how it is used. As professionals and recreationists generally submit reports at the end of the workday, Avalanche Forecasters use the prior day's reports to predict current and future conditions.

[7] https://parks.canada.ca/pn-np/mtn/securiteenmontagne-mountainsafety/avalanche/prevision-forecasting

Table 1 **List of participants, background information about them, and the studies they took part in. Participation is indicated with an "X"**

ID	Public Avalanche Forecasting Experience	Background	Public Forecasting Organization	Study 1	Study 2	Study 3	Study 4
1	N/A	Natural Science	Avalanche Canada	X			
2	N/A	Marketing, Communications	Avalanche Canada	X	X		X
3	4+	Geological Engineer	Avalanche Canada	X	X	X	X
4	10+	Mountain Guiding, Ski Patrol	Avalanche Canada	X	X	X	
5	5+	Mountain Guide, Educator	Avalanche Canada	X	X	X	
6	19+	Mountain Guiding	Avalanche Canada	X	X	X	
7	6+	Ski Patrol, Mountain Guiding	Avalanche Canada	X		X	
8	N/A	N/A	Avalanche Canada	X			
9	N/A	N/A	Avalanche Canada	X	X		
10	N/A	N/A	Avalanche Canada		X		
11	N/A	Engineering, Natural Science	Avalanche Canada		X		X
12	N/A	Mountain Guide, Communications	Avalanche Canada		X		
13	1+	Avalanche Field Technician	Avalanche Canada			X	
14	11+	Mountain Guiding, Engineering	Colorado Avalanche Information Center			X	
15	12+	Mountain Guiding, Geologist	Parks Canada			X	
16	19+	Geographer	Colorado Avalanche Information Center			X	
17	5+	Engineer	Avalanche Canada / DAC			X	
18	N/A	Educator	Avalanche Canada				X

Chapter 4. Research Approach and Methodology

4.1. Research Objectives and Questions

This work had several high-level research goals:

1. Understand and characterize the work and information processing practices, challenges, and needs in the domain of avalanche forecasting to identify tractable problems for visual analytics solutions.
2. Design and develop targeted visual analytics solutions.
3. Evaluate how designs support identified problems to inform further design.

These goals translated to the following research questions:

1. What are the work challenges, practices and needs in the domain of avalanche forecasting?
2. What representational and visual analytics design strategies might address the identified needs and challenges?
3. Finally, how do these approaches serve to address or impede the identified needs and challenges?
    a. Do current design principles in visual analytics apply well to these complex sensemaking domains?

4.2. Research Approach

The methodological approach to address these challenges is grounded in the tradition of problem-driven visualization research: the canonical visualization design study. Sedlmeir et al. (2012) provide a definition of a visualization design study:

> A design study is a project in which visualization researchers analyze a specific real-world problem faced by domain experts, design a visualization system that supports solving this problem, validate the design, and reflect about lessons learned in order to refine visualization design guidelines.

Given the multi-disciplinary nature of visualization as a field and the nature of designing visualizations for real-world problems, design studies draw on a variety of methodologies from other domains. These are qualitative studies aiming to develop understanding that may be transferred into other contexts with similar characteristics. To clarify the methods and aims in design studies, it is important distinguish between summative and formative research (Lam et al., 2012). Summative approaches aim to test the effectiveness of an interface, for example using laboratory-based studies for benchmark comparisons with metrics like completion time. Formative approaches aim to progressively and iteratively develop understanding to inform future design. Summative methods may sometimes be used in design studies, but the aims of such studies align more with formative methods. This is because summative studies require precise control of variables and factors which will not be known and will be difficult to anticipate in new problem areas. While the goal of summative research is generalizability and reproducibility, this is not the case in design studies. Here, the goal is *transferability* (Sedlmair et al., 2012) whereby the characterization of a problem space or solution space is sufficiently abstracted so that it may be potentially transferred to other contexts.

I now turn to a set of methodologies drawn on in this research and discuss their relevance to the aims of visualization studies and the present dissertation.

4.2.1. Quasi-ethnographic Methodologies

Design studies draw heavily on ethnographically-inspired methods to understand and characterize work practices (Carpendale, 2008; Lam et al., 2012; Munzner, 2009; Sedlmair et al., 2012). Traditional ethnography in anthropologies focuses on the study of culture. In engineering and visualization design, the focus is on technology and how it can best serve humans. It usually involves a long-term engagement to progressively develop a rich understanding of the target population, the data involved, analytic problems, and to abstract this into generic terms that lead to actionable design solutions. Ethnographic methods typically involve a mix of field observation and interviews, studying a target population in a naturalistic setting. As is well-known in the field of psychology (Nisbett & Wilson, 1977), introspection into past decision-making activities is unreliable and can only provide an incomplete picture. This is why the combination of observational methods and interviews yields far richer, more reliable, more useful insights than methods solely relying on retrospective introspection.

Problem characterizations produced through ethnographic methods are important contributions because they ensure shared understanding between researchers and domain experts, set design criteria, and establish the basis for evaluating visualizations and transferability (Munzner, 2009; Sedlmair et al., 2012). However, they remain relatively rare in visualization literature. This is not the case in more mature domains studying the use of technology in complex systems. Cognitive Systems Engineering (CSE) provides a robust methodological grounding for understanding “complex cognitive systems”, which are real-world work settings involving risk, uncertainty, ill-defined or competing goals, individual knowledge and reasoning, cognitive processes of groups and organizations, and interactions with computers and via computers (Hollnagel & Woods, 2005). As the avalanche industry apparatus and avalanche forecasting itself may described as a “complex cognitive system”, I draw on additional, ethnographically inspired methodologies from CSE in the following section.

Macrocognition and Cognitive Task Analysis

CSE researchers also argue for studies in naturalistic settings to understand the rich and contextual needs of domains where technologies are deployed (G. Klein et al., 2003). They typically distinguish between “microcognition”, addressing low-level cognition and perception studied through generalizable lab studies, and “macrocognition”, addressing higher-level cognitive processes that can only be studied in the real-world settings where they emerge and are used. This is similar to the conventional methodological framing of design studies in visualization, however, “macrocognition” is broader in scope because it considers all cognitive processes including those that might not directly relate to low-level visualization tasks but could and often do inform design.

This methodological lens is foundational to the field of cognitive systems engineering and has been applied in a variety of applied settings such as aviation, space shuttle mission control, healthcare, business, and control rooms among many others (Hollnagel & Woods, 2005). A common facet of this work is the use of cognitive task analyses (CTA) and cognitive process models to characterize a domain and guide design (Hollnagel & Woods, 2005). Cognitive task analysis provides a rich set of methods for studying cognition in real-world settings (Crandall et al., 2006). A range of methods varying in realism, task difficulty, generality, and temporal proximity (e.g., retrospective, current, or prospective) are applied in this work as needed.

4.2.2. Design Methodologies

After a foundational understanding of a problem domain has been established, iterative design and rapid prototyping begins (Munzner, 2009; Sedlmair et al., 2012). This involves cycles of design and feedback or evaluation with real users where each cycle further refines understanding of both the problem space, design strategies, or more technical system details. The “research through design” paradigm describes how the process of design and the artifacts produced through it may reflexively embody and provide knowledge about how it functions in a given problem domain (Zimmerman & Forlizzi, 2014).

Evaluation also plays a role throughout the iterative design process and may also be more summative or formative (Carpendale, 2008; Lam et al., 2012). For instance, long-term case studies of deployed systems employing multiple methods (Shneiderman & Plaisant, 2006) provide a rich multi-dimensional view of an artifact as it functions in the real-world, ensuring ecological validity. Again, ethnographically-inspired methods such as interviews and field observations are often used. Given the costs of developing fully functional prototypes, low-fidelity prototypes such as paper sketches or approaches that mimic the functionality of a system without implementing it, such as diary studies (J. Lazar et al., 2017) and wizard of oz studies (Dow et al., 2005), offer opportunities to evaluate alternative or successive design strategies and advance understanding. Other methods like think-aloud protocols (Cooke, 2010), involving participants vocalizing thoughts and reasoning during a task, or situated-recall methods (Bentley et al., 2005), where participants are probed using a video recording of their completing a task, can provide rich insights into thought processes associated with the use of technology without suffering the limitations of retrospective interviews.

Participatory Design

Often, stakeholders like domain experts are materially involved in the design process (Hartson & Pyla, 2019). This is often referred to as ‘participatory design’, ‘co-design’, or ‘co-operative design’. Having domain experts participate in design is advantageous in several ways. Through the creation of an artefact, participants may use and make apparent tacit knowledge that they are not aware of or cannot easily articulate. In addition, they may develop unique and creative solutions that designers or researchers might not have considered (Schuler & Namioka, 1993). This approach helps elicit core challenges, needs, and opportunities and ensures that visualization research leads to mutual benefits for researchers and collaborators (Jänicke et al., 2020). In doing so, it ensures the ecological validity and applicability of design artefacts. The co-creation of a shared artefact ensures that all parties are invested, relevant knowledge about the problems faced is elicited, and that solutions are more likely to successfully address the challenges faced.

4.3. Research Structure

This dissertation presents several studies addressing the stated objectives and research question and draws on the above methodologies. Study 1 (Chapter 5) uses quasi-ethnographic methods to understand and characterize problems in avalanche forecasting tractable for visual analytics. Study 2 (Chapter 6) progressively explores and refines visual analytics solutions to address the challenges of this problem domain and further refines the problem characterization through participatory design and interview methods. Study 3 (Chapter 7) employs a novel evaluation approach derived from CSE methodologies to provide a formative understanding of how existing and newly designed systems serve to meet the demands of this domain to inform future designs and to further advance problem understanding. Study 4 (Chapter 8) describes preliminary explorations of potential design solutions to address aspects of the problem domain identified in Study 1 but not further explored in other studies.

Chapter 5. Study 1: Understanding What Forecasters Do

> *It is the power of expectation rather than the power of conceptual knowledge that molds what we see in life not less than in art.*
>
> *(E.H. Gombrich, Art and Illusion)*

The goal of this preliminary investigation was to characterize the work and challenges and data practices of avalanche forecasters to inform potential visual analytics interventions and derive high-level design requirements. As is common when conducting visualization studies aimed at understanding work practices (Lam et al., 2012), I employed interview and observational methods as they provide rich data from which to gain a holistic understanding appropriate for this stage. I began with semi-structured interviews to understand how forecasters perceive and describe the challenges of their work (Part A). This broad view from the perspective of the forecasters helps identify the key issues forecasters face in their work and work practices. In Part B, I conducted field observations of forecasters on site, and, concurrently, video-recorded forecasters' workstations using these recordings as probes to more deeply explore the analytical reasoning involving the use of existing analytic tools. The understanding gleaned in Part A provided context for my investigation in Part B particularly during debrief interviews.

5.1. Part A. Forecaster Perspectives

The objective of this interview investigation was to understand how forecasters think about the challenges of their own work. Prior to the study, I understood that uncertainty and prior knowledge play an important role in avalanche forecasting, but wanted to understand how forecasters think about these issues and their relationship to data analysis.

5.1.1. Research Questions

1. How do forecasters use and think about the data available to them?
2. What are the key challenges forecasters face in their work?
3. What role do knowledge and experience play in analysis?

4. How do forecasters think and reason about uncertainty?

5.1.2. Method

Semi-structured interviews with 5 professional avalanche forecasters were conducted on Avalanche Canada premises in Revelstoke, British Columbia during the forecasting season. Participants were asked about common work practices and challenges in avalanche forecasting, the role of data and evidence, the role of prior and tacit knowledge, issues of collaboration, and issues of uncertainty. Interview questions began open-ended to avoid leading participants: "Can you walk me through a typical forecasting day?", "What are the biggest challenges in your work?", or "What are some common uncertainties you deal with?". Follow-up questions about topics of interest were asked to refine understanding. The interviews were audio-recorded and then transcribed.

5.1.3. Participants

Table 2 Participant table for Part A of Study 1

ID	Public Avalanche Forecasting Experience	Background
1	N/A	Natural Science
2	N/A	Marketing, Communications
3	4+	Geological Engineer
4	10+	Mountain Guiding, Ski Patrol
5	5+	Mountain Guide, Educator

5.1.4. Analysis

Data were analyzed using thematic analysis (Braun & Clarke, 2012). Transcripts were concurrently segmented (Geisler & Swarts, 2019) and coded according to emergent themes by one coder. The codes were then refined in two passes. These themes were then grouped into thematic categories (Table 3). Inter-rater reliability was measured with one other coder who had a background in avalanche research and limited experience in qualitative research methods using a transcript sample representing 10 percent of all data (Geisler & Swarts, 2019). Simple agreement for high-level themes was 0.89, Cohen’s Kappa was 0.81, and Krippendorff’s Alpha was 0.82.

For the sub-themes, simple agreement was 0.75, Cohen's Kappa was 0.70, and Krippendorff's Alpha was 0.71. All analysis was conducted using Microsoft Excel[8].

Table 3 Table of main themes, sub-themes, their definitions and their relation to the broader topics of data, analytic process, and collaboration & communication in ambiguity.

	Main Theme	Sub-Theme	Definition
Data	Missing Information	Explicit	Missing information is explicitly represented in data.
		Implicit	Missing information must be inferred from the situational context.
	Data Representativeness	Classification Overlap	Classifications are often not independent or mutually exclusive.
		Conservative Bias	Avalanche professionals are conservative when faced with uncertainty in the field or in data.
		Circumstantial Definitions	Official definitions and unofficial practices for reporting and interpretating data depend on the situational context.
Analytic Process	Analytic Practices	Subjective Hunches	Considering the behaviour, concerns, and hunches of others to inform and guide analysis and interpretation.
		Immersion	Forecasters spend several days forming a mental model through an *undirected* review of contextual information.
		Context-seeking	*Directed* information search for supplementary contextual information.
		Mental Projection	Forecasters assimilate information by imagining and mentally visualizing the interactions of the snowpack, weather, terrain, and people.
		Updating	Forecasters iterate over knowledge artifacts like their forecasts as they conduct their analysis and update their own mental models.
		Deliberate Omission	Forecasters manage information overload by deliberately ignoring some information.
	Analytic Challenges	Lack of Good Representations	Forecasters lament a lack of good visual representations to alleviate some cognitive effort.
		Lowering Danger Ratings	It is challenging for forecasters to lower danger ratings as data reveal instability rather than stability

[8] https://www.microsoft.com/en-ca/microsoft-365/excel

	Main Theme	Sub-Theme	Definition
Collaboration & Communication		Continuity	Forecasting relies on the continuity of analysis and monitoring. Shift changes disrupt this continuity.
		Translating Analysis	Forecasters struggle with communicating complex conditions with simple clarity.
	Collaborative Sensemaking Strategies	Data Production	Forecasters facilitate collaborative work by producing hand-off notes and other internal documentation.
		Regular Discussion	Forecasters draw on each other's diverse knowledge through daily discussions.
		Reaching out Directly	Forecasters call or email field operators for further information when faced with critical information gaps.
		Professional Exchanges	Forecasters work on-site at other agencies to gain a deeper understanding of how data are produced and what they mean.

5.1.5. Findings

In the following section, I present findings from the interview studies using **bolding** to highlight key themes that emerged. Findings are organized by topics related to data, analytic process, and collaboration & communication. A comprehensive set of themes and definitions may be found in Table 3 at the end of this section.

Data Challenges and Practices

Forecasters informed me that the data used in avalanche forecasting are uncertain, have ambiguous expressions or meanings, and have biases. These characteristics lead to ambiguity and a need to consider alternative interpretations beyond what is explicit in data.

Forecasters said one of their key challenges is the uncertainty involved in data sparsity or **missingness**. Data are often **explicitly** missing as is the case when remote sensors malfunction or fail to transmit.

> [Weather stations] that have good weather or wind information are even less, and then that's if they're even reporting… (P4)

Missingness might also be **implicit**, having to be inferred from the given situational context.

> In a large storm that closes highways and grounds helicopters, it's very common the next day to not get any avalanche observations... but the weather and your personal experience very much suggests that there was going to be an avalanche cycle... (P1)

Forecasters rely on contextual information to understand how to appropriately interpret data following **circumstantial definitions**. Some of these contingencies are officially documented or ingrained within formal procedures, while others are only learned through extensive experience and knowledge.

> The… courses do quite a good job of standardizing those kinds of threshold amounts [but] people who have spent a lot of time on the coast [...] may think a 30 centimeter storm doesn't really do very much... (P1)

Common to many classifications of the complex natural world, avalanche **classifications overlap** and are not mutually exclusive. Technically accurate hazard assessments might include several overlapping avalanche types resulting in overly complex public communications. Instead, forecasters try to choose a subset of avalanche types based on what may inform optimal risk mitigation strategies by the public.

> When you're modeling the natural world, you take shortcuts and there's simplifications... they don't occupy fully independent places... we sometimes have to have discussions about whether we want to be technically accurate, or whether we want to retain clarity... that starts to get quite complicated... we look for ways to simplify... (P1)

The nuances of evidential reasoning and interpretation of data in avalanche forecasting also extend to the risk-based **conservative bias** some operators may have. This has particular bearing when interpreting other's hazard assessments such as those provided in the InfoEx. Some may be more or less conservative, and forecasters have to factor in such considerations when weighing evidence.

> [Discussing varying risk tolerance] Another forecaster would have said something like: '... they always call that a little more than what it actually is.' ...that may influence me to say: 'Okay, well, maybe I should not necessarily discredit it, but I put less weight into it...' (P3)

Analytic Processes and Reasoning

Forecasters employ a variety of sensemaking strategies involving speculation and imagination. They integrate their prior knowledge, experiences, and contextual clues in data to synthesize understanding and explore risk implications.

Forecasters synthesize, evaluate, and integrate information using a mental simulation technique they call **mental projection**. It is a process of imagining oneself in the field to understand conditions and their risk implications.

> [T]hat's a technique that a lot of people use to help forecast... kind of projecting yourself mentally, whether you close your eyes or you just have some kind of image of the kind of slopes, the kind of areas where the people are moving around... I think that experiential part there is really relevant to the process... (P1)

This might involve **mentally converting biases** such as wind data from weather stations in windy locations.

> [T]here can actually not be that much wind in the park and you can have 60 kilometers an hour winds at that station... taking an input and then adjusting it for myself... (P2)

It might also involve simulating alternative future scenarios and their risk implications.

> If things are a little bit unusual, I... try and strip it down and build some kind of synthetic profile either in my mind, or sometimes even do it on the whiteboard... and then figure out the most likely, it's usually a set of scenarios... (P1)

Forecasters describe their work as Bayesian-like because they are constantly updating their mental models with new information and **deliberately omitting** weak or redundant evidence. They reported having to **immerse** themselves in data over several days of their shift to build confidence in their sense of understanding. This often involves undirected explorations of general background information.

> ...a day, you know, more likely two days to become fully sort of understanding of what's going on in your region... even if you can read it all in a day, it takes a little time for it to sort of percolate and for you to understand what that means... (P1)

To address identified gaps in understanding, forecasters actively **seek contextual sources** of information.

> I'll... look for keywords like 'oh ya... skiing, like, steep terrain in the Alpine, up to 40 degrees and just exposed features. No problem.' That tells me that not much is going on. Yeah, people are confident... (P2)

As they conduct their assessments, they iteratively **update** knowledge artifacts like the public bulletin to match their current understanding.

> I’m pretty iteratively making small changes in the forecast... I'll just move that right into the forecasts, put it there, save, and I go back to what I was doing... (P2)

Collaborative Challenges and Practices

Collaboration helps individual forecasters overcome the limitations of their own knowledge by drawing on the collective knowledge and experiences of their peers. At the same time, communicating the complexity of their assessments in simple terms is a constant challenge that creates ambiguities.

Forecasters vary in knowledge and experience which likely contributes to some variations in interpretation. However, this diversity is seen as an advantage as, collectively, it addresses the gaps in understanding any single forecaster may have.

> My experience may be different from you know... another forecaster’s experience and I can learn from that person... there's those kinds of exchanges that happen... (P1)

Forecasters share knowledge and solicit their peers' perspectives in daily **discussions**.

> At two o'clock, we have our pow-wow where we all kind of go through our hazards and our problems... it's kind of like a peer review session... (P3)

Professional exchanges with partnering operations help avalanche forecasters enrich their understanding of how data are produced in a variety of operational contexts.

> Whether that's highways or ski hill, snowcat skiing, heli-skiing... there's variability between the individual operators... And the only way to really fully understand is to go and spend a bit of time with that operator... We have professional exchanges go on... (P1)

Forecasters also phone operators and **reach out directly** for clarification or if they are uncertain about how they should be thinking about conditions.

> If I am potentially missing something or I just don't feel comfortable... I'll start picking the phone up and trying to find people in the area that can provide more, more insight... (P3)

Collaboration allows forecasters to account for each other's knowledge gaps, at the same time, it presents challenges such as communication of analysis. Forecasting relies on the **continuity of analysis**. Shift changes can disrupt this continuity and forecasters struggle with communicating relevant details as part of the hand-off process.

> There's a lot of variability in different people and... what sort of information they leave... that's the first place I'll look... hoping that the... previous forecaster has left enough information to start that picture... (P3)

To facilitate the hand-off process, forecasters **produce knowledge artifacts** like dedicated hand-off notes or detailed descriptions of snowpack stratigraphy.

> Talking about hand-off notes, I am trying to take that ease and control that I have at day four or five... and I give that to the next person, so they don't feel like they have to do their process of discovery from ground zero essentially... (P2)

Forecasters told me that this is seen as a separate and additional task often completed at the end of the day when forecasters are fatigued. This is why documentation used in support of hand-off and collaboration is often incomplete. Whether communicating to fellow forecasters or the public, capturing complexity and nuance in simple and understandable terms is a challenge.

> To simplify it... that's when you are kind of having to use your own best judgment... (P2)

Forecasters face additional challenges when communicating their understanding to the public. Forecasters must **translate** their understanding and cater it to an audience that varies in understanding and expertise. This often involves exploring alternative future scenarios, their implications, how an audience may interpret what the forecaster is saying, and subsequently choosing a simple communication strategy that comprehensively accounts for these alternatives.

> So instead of trying to write my forecasts like: 'oh, if we get 10 centimeters it will probably be okay, but if we get 20, then it'll probably come unglued'. It's like 'just watch for conditions to change as you increase with elevation... if it starts to feel stiff or slabby underneath your feet... use that terrain feature to go around it.' (P2)

5.2. Part B. Observing Foresting Analytics

Several weeks after having conducted interviews and analyzed the interview data, I returned to observe forecasters in the workplace and use this as a basis for more targeted inquiry and discussion. The purpose of this portion of the study was to refine my understanding of forecaster's work practices and observe how tools and data are used in practice. In addition, this provided an opportunity to compare what forecasters say and what they do.

5.2.1. Research Questions

1. What types of tools do forecasters use? How do they use them?
2. How do forecasters think about and use data?
3. What is the role of prior knowledge and experience?
4. How do forecasters collaborate?

5.2.2. Method

Part B focused more on understanding how current tools are used. I conducted field observations on Avalanche Canada premises for a week, collecting field notes and audio recordings of daily discussions. Observations are an important method for evaluating work practices (Lam et al., 2012) as they preserve realism and the broader working context. Thus, they address the limitations presented by prior interviews where data only reflect the perspectives and accounts of forecasters.

At the same time, I employed a method called cued-recall debrief (CRD), a situated recall method developed for use in complex decision-making contexts (Omodei & McLennan, 1994) and adapted for human-computer interaction (Bentley et al., 2005). This method was chosen because it improves reliability and recall of prior decision-making behaviour without being disruptive. While methods like think-aloud protocols offer similar benefits, vocalizing thoughts and completing a task concurrently introduces additional cognitive load and changes the nature of the task. Moreover, this would be disruptive to others in a shared office environment. Debriefing forecasters using screen recordings of their workday provides a contextualized shared point of reference from which to make inquiries about how data and analytics tools are used, the reasoning processes involved, and how they fit in the broader context of work.

Seven forecasters were observed in the field and 4 were debriefed using CRD. Camcorders positioned behind workstations in view of monitors and the desk surface captured recordings of forecaster's workday and their use of technology as well as artifacts such as hand-written notes (Fig. 6). Throughout the workday at convenient time intervals like lunch breaks, I removed memory cards from camcorders and reviewed recordings at another location. I noted timestamps in video recordings when I observed a behaviour that I did not understand or was of interest. For instance, when forecasters were viewing charts, switching between tools, scrolling or mousing over data tables, and other similar behaviours involving tool use, data, and analytic reasoning. Debrief interviews occurred on average one hour after the end of the workday to allow for a break. During the debrief interviews, recordings were played back to forecasters at marked timestamps, and forecasters were asked to explain their thought processes and actions. I asked questions like: "Can you explain what you were doing and thinking here?" These debrief interviews were video recorded and transcribed.

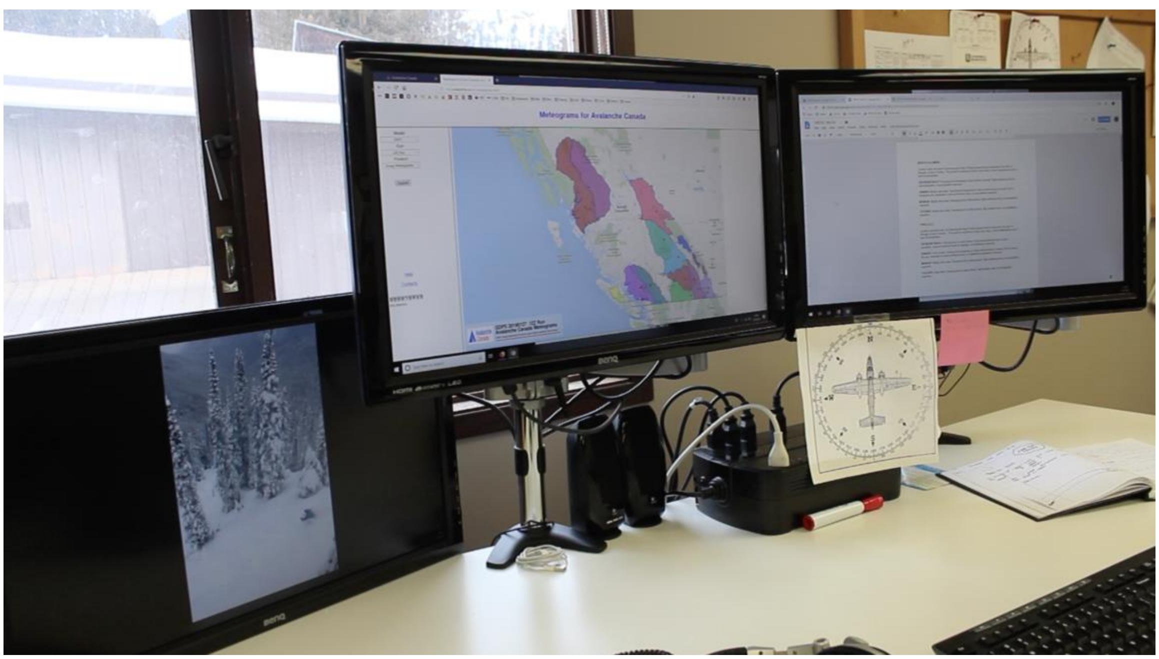


Figure 6 A sample screenshot of a forecaster workstation recording.

5.2.3. Participants

Table 4 Participant table for Part B of Study 1

ID	Public Avalanche Forecasting Experience	Background	Data Gathered (O = Field Observation, CRD = Cued-Recall Debrief)
2	N/A	Marketing, Communications	O
3	4+	Geological Engineer	CRD / O
4	10+	Mountain Guiding, Ski Patrol	CRD / O
5	5+	Mountain Guide, Educator	CRD / O
6	19+	Mountain Guiding	CRD / O
7	6+	Ski Patrol, Mountain Guiding	O
8	N/A	N/A	O

5.2.4. Analysis

I applied the thematic coding scheme developed in Part A (Table 3) to notes and transcripts in Part B (Table 5). Thematic coding was applied by one coder in two passes.

5.2.5. Findings

In the following section, I present findings from the interview studies using **bolding** to highlight key themes that emerged. Findings are organized by topics related to data, analytic process, and collaboration & communication. A comprehensive table with evidence gathered as well as the method it was sourced with is available in Table 5 at the end of this section.

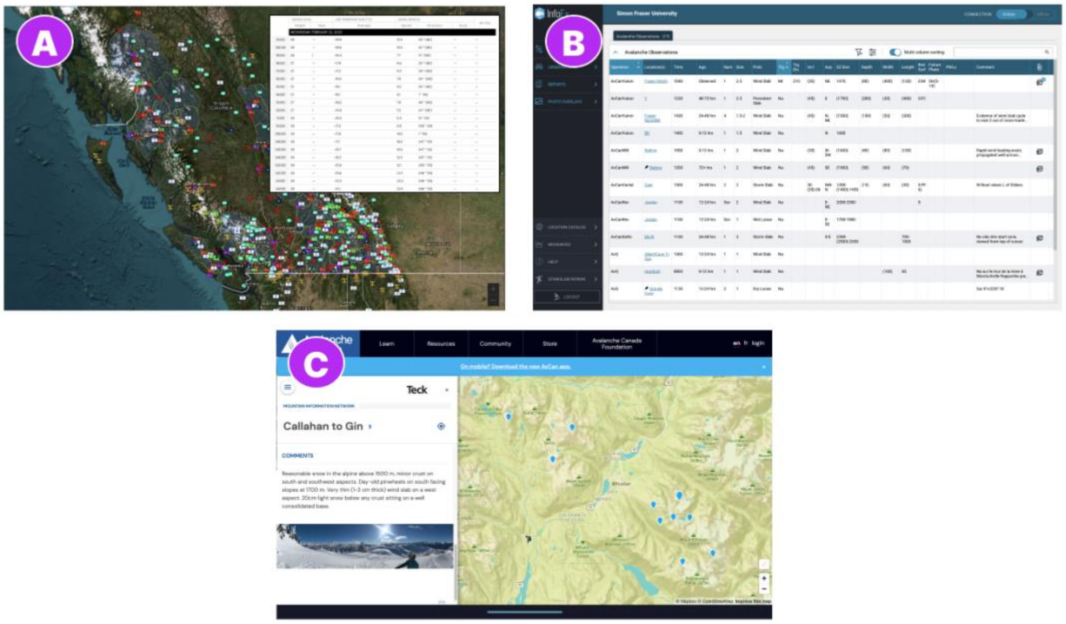


Figure 7 **The existing fragmented analytics systems used in forecasting. A) The Avalanche Research Forecasting Interface (ARFI) organizes various forecasting resources such as weather stations in a geospatial format. B) The professional Information Exchange (InfoEx) where avalanche professionals share field observations. C) The Mountain Information Network (MIN) where recreationists share reports.**

Analytic Workflows and Tooling

I found that forecasters work individually and collaborate synchronously as well as asynchronously and sequentially on their daily bulletins. Forecasters generally start the day reviewing forecasted weather using numerical weather prediction models, covering large continuous spatial areas, and then analyzing point data of weather, snowpack structure, and avalanche conditions reported by avalanche safety operations contributing to the InfoEx and automated weather stations among several other data sources. A large portion of the morning is spent investigating numerical weather prediction models as well as meeting with professional meteorologists who provide additional guidance on model interpretation. The tools and workflows used for meteorological models were well-integrated and smooth. However, I found the tools used to investigate human and weather station observations were visually and procedurally fragmented creating friction in accessing and synthesizing these data (Fig. 7). In one instance, a forecaster opened several browser windows of data tables to investigate temporal trends while in several others, forecasters were navigating between several windows to while make comparisons about data or switch between spatial, temporal, or informational views of the same data. This fragmentation is one of the motivating factors behind this research collaboration with Avalanche Canada. The aim is to develop visualizations that aggregate these disparate data in a centralized platform.

Previous efforts to consolidate these resources led to the creation of a map-based portal (Fig. 7.A), which spatially organizes hyperlinks to access primarily meteorological data from remote sensing weather stations and cameras. Using this tool, forecasters must open multiple browsers for different websites to view weather station telemetry from individual weather stations and primarily in a data table format.

Data tables also played an important role in the investigation of human field observations. Forecasters used two platforms for investigating these reports: the Canadian Avalanche Association's Industry Information Exchange (InfoEx) (Fig. 7.B) for avalanche professionals and the Mountain Information Network (MIN) (Fig. 7.C) for recreationists. Throughout these tools, we observed forecasters making extensive use of data tables. They would visually scan these tables, noting trends, central tendencies, and other patterns while at the same time, reading unstructured text from report authors to glean enough context so as to understand their significance. Although simple visualizations were available in the professional InfoEx system, we were surprised to find that they were rarely used.

Talking About Data

I found several examples showing how organizational knowledge relevant to the nuanced interpretation of data is in large part oral tradition exchanged through the shared practice and environment of work.

I observed several **discussions** that dealt with the topic of how to interpret reports from the field. For instance, one discussion dealt with the interpretation of a report authored by an operator who was known to have a **conservative bias** and what the implications of this were for hazard assessments. In another discussion observed in the field, a junior forecaster with a guiding background described how they are coming to understand the challenges of their new remote work environment, noting the nature of what types of information may be missing.

> After having worked this job [Avalanche Canada] ... I sort of realize the big holes the operators leave in their writeups [...] because they are having face-to-face conversations... and maybe not putting that information into their writeup... saying this layer [of snow] does not exist in our area may not be helpful to them, but it really helps us here in this office... (P8)

How **classifications** and **circumstantial definitions** are applied in hazard assessment and risk communication was also a frequent topic of conversation.

> I like [X's] point yesterday, wind slabs in the alpine are kind of like cornices that you find always... it is just a winter mountain hazard... it goes on the bulletin when it is elevated to more than normal caution... (P2)

Tacit Sensemaking and Analytic Processes

Early sensemaking processes, particularly those involving personal experiences or trust, may be difficult to articulate out of context and consequently, share with others.

When debriefing forecasters about their workday I found they relied on the **subjective hunches** of operators that they personally trusted and were more familiar with. This factored into how evidence was weighed, and the confidence forecasters had in it.

> I feel good about who was about in the operation. So, I felt that the test was valid and valid information that I should be thinking about... (P3)

I also found forecasters exploring general contextual information to **immerse** themselves. They found it difficult to articulate how they were using the information, reflecting characteristics of early sensemaking processes.

> It was just to give me an orientation to get my mental picture for forecasting... just a little bit of context... I don't know what that does for me exactly... (P4)

Collaboration and Knowledge Artifacts

The bulletin serves as a knowledge artifact representing a forecaster's current understanding of avalanche conditions. The bulletin scaffolds analysis and guides information search, particularly during hand-off at shift changes. However, the reasons behind specific changes to the bulletin are not always explicitly captured leaving future collaborating forecasters to speculate about the reasoning that might have been involved.

Forecasters do not just iterate over their own bulletin over the course of the day, they often carry forward the previous day's bulletin even if another forecaster wrote it. I observed how forecasters **update** it as they formulate their own new understanding.

> I import yesterday's forecast... and I tweak my forecast so it matches my nowcast... (P6)

The specific reasons behind these updates are not made explicit, leaving the forecasters coming on shift to **seek contextual** information to speculatively reconstruct their coworker's evidential reasoning process.

> So I reviewed a few avalanches to understand what was driving those avalanches and why [X] added that persistent slab problem again... (P6)

Table 5 Table of evidence gathered in Part B corresponding to themes identified in Part A.

	Part A Main Theme	Part A Sub-Theme	Part B Evidence (O = Observation, CRD = Cued Recall Debrief)
Data	Missing Information	Explicit	
		Implicit	O
	Data Representativeness	Classification Overlap	O
		Conservative Bias	O
		Circumstantial Definitions	O
Analytic Process	Analytic Practices	Subjective Hunches	CRD
		Immersion	CRD
		Context-seeking	CRD
		Mental Projection	
		Updating	CRD
		Deliberate Omission	CRD
	Analytic Challenges	Lack of Good Representations	CRD
		Lowering Danger Ratings	
Co		Continuity	O

		Translating Analysis	O
	Collaborative Sensemaking Strategies	Data Production	O
		Regular Discussion	O
		Reaching out Directly	O
		Professional Exchanges	

5.3. Discussion

The described findings demonstrate how avalanche forecasting involves knowledge and interpretation that goes beyond a face-value read of data. Considering and evaluating alternative interpretations based on prior knowledge is a pervasive facet of their work. This study surfaced such issues which I characterize and abstract as dealing with ambiguity. I consider three areas where ambiguity arises: data, analytic process, and collaboration and communication. I draw this distinction to abstract and direct attention toward general issues of ambiguity. At the same time, I acknowledge these are categories are not entirely independent.

5.3.1. Data

Ambiguity emerges from data because they are incomplete simplifications of the complex phenomena they represent. Ambiguity may be present in the expression of data or in how representative data are of the phenomena of interest. Forecasters use their knowledge, experience, and cues within the data to explore plausible explanations that account for what they see. This might involve speculating about causal factors explaining the shape of data, imagining how all the known information can be synthesized into a coherent understanding, or extrapolating point data across space and time. The challenges of data render the consideration of alternative perspectives a core and functional aspect of forecasters' work. This suggests that visualization and representational design approaches should carefully consider when and where this form of sensemaking is relevant and ensure that it is not impeded.

5.3.2. Analytic Process

Ambiguity plays a productive role in the analytic process but also presents challenges in managing and navigating analyses. This involves distinct cognitive and analytic processes that consider alternative interpretations and involve the active manipulation and construction of knowledge. For instance, forecasters extrapolate alternative future scenarios in risk analysis and risk prediction. Through mental visualization or explicit sketches, forecasters explore and evaluate alternative future scenarios to make predictions, explain data, identify risk implications, and choose ways to best communicate this understanding to others. In one example, forecasters described drawing diagrams of snowpack structure to explore how it may evolve and react to different potential future weather scenarios. These processes could more explicitly be supported in visual analytics systems. For instance, physical snowpack simulation, which are increasingly being used to supplement understanding of existing snowpack conditions (Morin et al., 2020), could be extended in interactive visualizations that allow forecasters to explore alternative future conditions and how the snowpack could react.

These analytic processes also present challenges for navigating one’s own analysis. The judgments and analytic choices made present alternative potential analytic paths through data. As forecasters weigh evidence and update their understanding of avalanche conditions, they iteratively adjust knowledge artifacts to match their understanding. However, how specific evidence was treated and led to certain assessments is not explicit, is often forgotten, and may be difficult for forecasters to reconstruct. I conjecture that this can make it difficult for forecasters to coordinate their own work as well as articulate their own analyses to others. This suggests mechanisms that can make past analytic steps and reasoning processes apparent could help forecasters manage and coordinate their own analyses. By externalizing the relevant elements to consider in assessments and making their relations apparent, an external representation can reduce the cognitive costs of reasoning about evidence, coordinating or planning work, and navigating a problem space (Kirsh, 2010). Visual analytics methods for analytic provenance (K. Xu et al., 2015) and interaction traces (Vuillemot et al., 2016) be of benefit here.

5.3.3. Collaboration and Communication

Ambiguity also arises from the complexities of collaboration and communication. Each forecaster holds a unique perspective and interpretive lens, presenting a form of ambiguity. Forecasters use strategies such as regular discussions or hand-off notes to exchange knowledge and clarify contextual evidence, explain reasoning, and to provide context that served as the basis for assessment decisions. This serves to enrich the set of potential interpretations to consider and can help disambiguate meaning by clarifying the most plausible explanations and interpretations of data. However, due to the effort required for this task and the difficulty in anticipating what may be relevant, such information is often incomplete, leaving forecasters to speculate about their colleagues' reasoning processes. In addition, the forecasters' organizational knowledge is predominantly oral tradition exchanged in application to the immediate demands of work, making it vulnerable to being lost and leading to redundant discussions when training new staff. This suggests that lightweight mechanisms, integrating with existing workflows and practices, that allow key evidence and relevant knowledge to be captured during analysis could be helpful here.

Forecasters simplify their complex understanding of avalanche conditions to ensure that members of the public, whether novice or expert, can apply appropriate risk-management strategies. In doing so, forecasters mitigate the risks of potential scenarios the public might encounter or the confusion that might result from overly technical communications. Often, this involves reconciling alternatives based on risk. For instance, when two avalanche problem types require the same risk mitigation strategies, forecasters will use one of them and supplement any further guidance using plain and actionable language. The myriad ways to communicate hazards presents its own form of ambiguity. The contextual factors involved in these decisions are not explicitly captured making it difficult for the organization to a.) describe how and why these factors lead to certain assessment decisions and b.) define how these factors should be used in assessment decisions. Tools to capture the use of knowledge and how it relates to evidence stake to not only benefit day-to-day operational collaborations but also provide the basis for developing further organizational knowledge and procedures. A corpus of meta-data pertaining to how evidence is used in assessments could serve to highlight the contextual factors involved, how they are used, and set the basis for investigations of how such factors should be used in assessment decisions.

5.4. Limitations

It is important to note that while Part A of this study involved additional coders to assess reliability, the data from subsequent parts were analyzed solely by me. The use of multiple methods including interviews, observations, and situated recall in this study strengthens the validity and reliability of the findings. However, analyzing data only by myself presents limitations in reliability and validity. This serves as a motivation for further research, which is necessary to refine the understanding of the avalanche forecasting domain, particularly in relation to data, and to explore the potential of visual analytics as a supportive tool.

Chapter 6. Study 2: Designing Visual Analytics Solutions

I present a participatory design study with Avalanche Canada forecasters to develop two visualization prototypes, WxObs and AvObs, addressing the challenges of ambiguity dealing with data. This study begins the stage of iterative design and rapid prototyping through which understanding of data and tasks is further refined and abstracted in more general terms to support targeted design solutions. By designing artefacts in collaboration with experts and studying its application in operations, lessons about the nuances of data, their use in analysis, and the potential for visualization to support these data and analytic processes is better understood. While the interviews, observations, and situated recall in the prior study offer some insight into this, understanding remains coarse. The objective of the present study is to refine this problem characterization of the data and task within the avalanche forecasting problem domain and to begin the investigation of how visualizations may serve to offer support for the associated demands.

The two prototypes and their respective data sources were chosen as they are of a relatively higher priority to avalanche forecasters than others. WxObs aggregates weather station telemetry for weather forecast validation and real-time monitoring as weather systems evolve. Meanwhile, AvObs concentrates on human field-reported avalanche observations, which are the strongest form of observation (Class 1) as they are the most direct evidence of structural instabilities in the snow. The design process, design features, and findings from forecasters and my own reflections are discussed for each prototype.

6.1. Research Questions

1. How can visual analytics tools be designed to address the challenges of ambiguity in avalanche forecasting, particularly when dealing with weather observations or avalanche observations? How can visualization designs **provoke** or otherwise facilitate ambiguous sensemaking?
2. How do the developed visualization prototypes (WxObs and AvObs) support or impede the forecaster's ability to interpret and analyze data and how can these

insights inform the further design iterations pursuant towards supporting ambiguity?

6.2. Method

As is common practice at this stage of a visualization design study (Munzner, 2009; Sedlmair et al., 2012), prototypes were iteratively refined from a broad set of alternative designs in the form of paper mockups to interactive visualizations with live operational data over the course of a summer (off-season) and winter forecasting season (Fig. 8). During design, I informally employed a variety of evaluation methods such as think-aloud protocols and unstructured interviews to better understand how tools are used. In addition, forecasters themselves provided visual sketches and contributed design ideas. Data collected include design artefacts such as prototype sketches and the prototypes themselves, written documentation, and email exchanges, as well as video and audio recordings from discussions and think-aloud sessions.

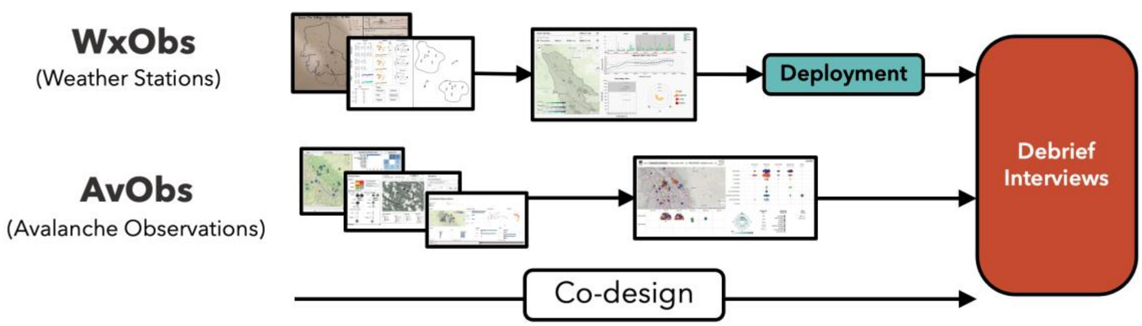


Figure 8 Process diagram describing iterative participatory design approach for both WxObs and AvObs prototypes.

During the study period, WxObs was used operationally for the latter part of the winter season. A retrospective interview was used to solicit feedback from forecasters and understand how the tool was being used in practice. At the same time, AvObs was still being designed and developed. Due to these differences, findings for each prototype tool are presented in a different format. The design and findings of WxObs are separated because I report on lessons learned from retrospective interviews about a real-world field deployment of the tool. By contrast, as AvObs involved an iterative design process, involving key findings at intermediary stages informing the design of subsequent iterations, a description of the design artefact and design process is included as part of the findings. This aligns with reporting practices in “research through design” (Gaver, 2012)

6.3. Participants

Seven forecasters participated in the design of both prototypes (P2-P6 ; P11, P12) while two others (P9, P10) only provided feedback after having used WxObs in practice. In addition, my supervisor and I were directly involved in the design process.

Table 6 Participant table for Study 2

ID	Public Avalanche Forecasting Experience	Background
2	N/A	Marketing, Communications
3	4+	Geological Engineer
4	10+	Mountain Guiding, Ski Patrol
5	5+	Mountain Guide, Educator
6	19+	Mountain Guiding
7	6+	Ski Patrol, Mountain Guiding
10	N/A	N/A
11	N/A	Engineering, Natural Science
12	N/A	Mountain Guide, Communications

6.4. Analysis

Analysis followed a combination of thematic analysis (Braun & Clarke, 2012) and my reflections of the design process (Zimmerman & Forlizzi, 2014). My analysis is informed by my direct involvement in the design process and is therefore a reflection of the understanding I formed engaging in this participatory design study. I reviewed data and artefacts collected throughout this period and drew on my first-hand experience to identify themes, quotes, and evidence that aligned with the key insights we had throughout the design process. I organized evidence according to a set of themes that best reflected my understanding of the important aspects of the design process. This process was iterative, involving several passes to extract themes as well as discussions with all participants to confirm common understanding.

6.5. WxObs

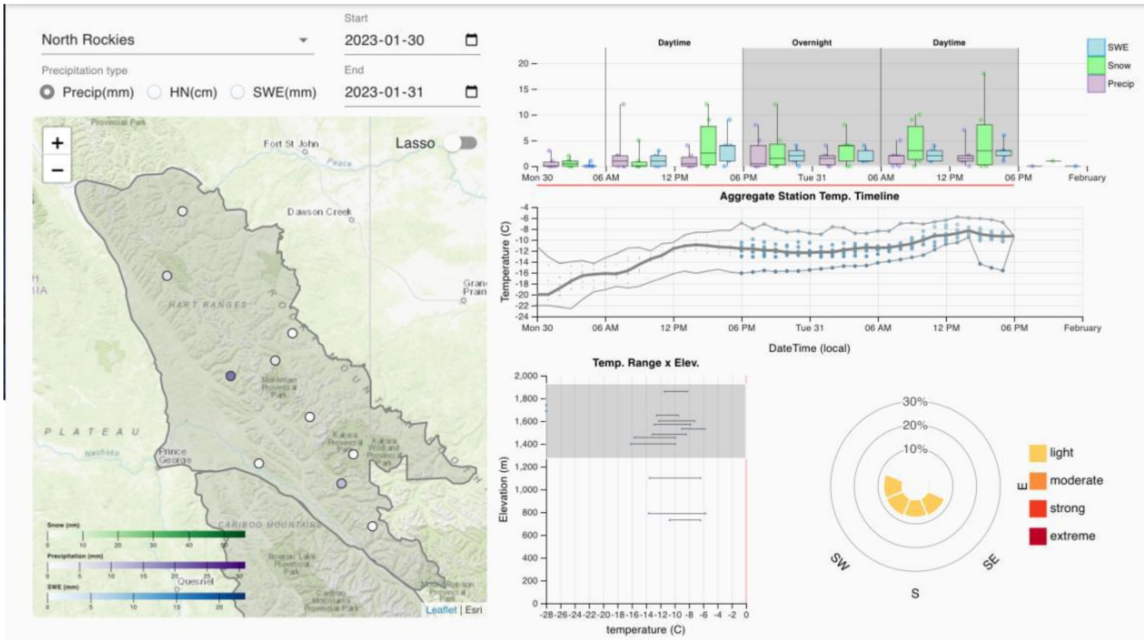


Figure 9 The WxObs prototype for analyzing weather station telemetry.

6.5.1. Design

The involved forecasters and I employed a conventional design approach for a multi-view coordinated visualization (Fig. 9). Our design intent was to address the fragmentation of navigating disparate resources through a unified representation offering visual summaries and explorability. Our design approach follows the information-seeking mantra “Overview first, zoom and filter, then details on demand" (Shneiderman, 1996). Numerical aggregations for each telemetry type (wind, temperature, and various measures of precipitation such snow water equivalent, snow height, or liquid precipitation), offer an initial summary “overview". “Multiple levels of detail" are shown in various visualizations which are linked and coordinated with “brushing", “zooming", and “filtering". A tooltip for individual weather stations shown as points on the map show all telemetry at that location for “details-on-demand".

6.5.2. Findings

The ambiguity of uncertainty

The operational use of the WxObs prototype highlighted how analysis of weather station telemetry presents issues of **data uncertainty** that give rise to ambiguity. There are too few weather stations to capture the variability of weather conditions in the mountains (Lundquist et al., 2019) and the weather stations that do exist, produce data that is often affected by multiple potential sensor and transmission errors (Mekis et al., 2018). Discerning potential transmission errors and how individual weather stations come to represent broader weather patterns is a matter handled through the forecaster's judgment and interpretation. Forecasters normally use text tables to view each weather station's telemetry individually and progressively build up an understanding of weather patterns.

This bottom-up approach stands in contrast to our top-down and overview-first visualization designs. Aggregate and visual summary measures and linked interactions introduced a **new and unfamiliar approach** that challenged forecasters.

> I've always looked at the data in a pretty disaggregated way... What I'm having to learn is to kind of let go of that, needing to see the disaggregated view first so that I can aggregate the data in my brain so to speak... (P12)

The need for raw data

Forecasters kept returning to the **tabular format** to see **raw data**. This went beyond simply being about path dependence and familiarity with this visual format. Forecasters could not **trust** numerically aggregated results. They have developed visual scanning strategies to detect errors in the data.

> It largely stems from the trustworthiness of the data… I like things in their raw format just for my own sake… my own stamp of approval… I guess it's easy for my eyes to decode differences or irregularities. (P3)

Forecasters also returned to this raw data table form as a way to better understand how visualizations were manipulating data and to **scaffold the learning** of the analytic affordances of these new tools.

> having [raw data table] side by side with the visualization helped me to interpret: Okay, what's the visualization trying to tell me here? (P4)

Forecaster reflections

Forecasters who adopted the WxObs visualizations more readily in their work found the tool provided them with a richer and deeper understanding of meteorological phenomena than traditional data tables alone. Drawing a historical comparison to the role of computers in meteorology, forecasters view visualizations as a steppingstone in a **transitional phase** towards more data-driven and modeling-based approaches.

> There was a transitional phase there where the computer was more an aid to help the forecaster make some initial assumptions... then the forecaster would tweak the forecast and actually write the forecast manually still... and now we're to the point where that really isn't the case... (P12)

6.6. AvObs

6.6.1. Findings

As I am reporting on a process of collaboratively designing the AvObs product, the design and the feedback and processes that led to this design are interleaved and presented in such a way that the design artifact, the design choices involved, and what led to those choices are treated as research findings.

Initial Design

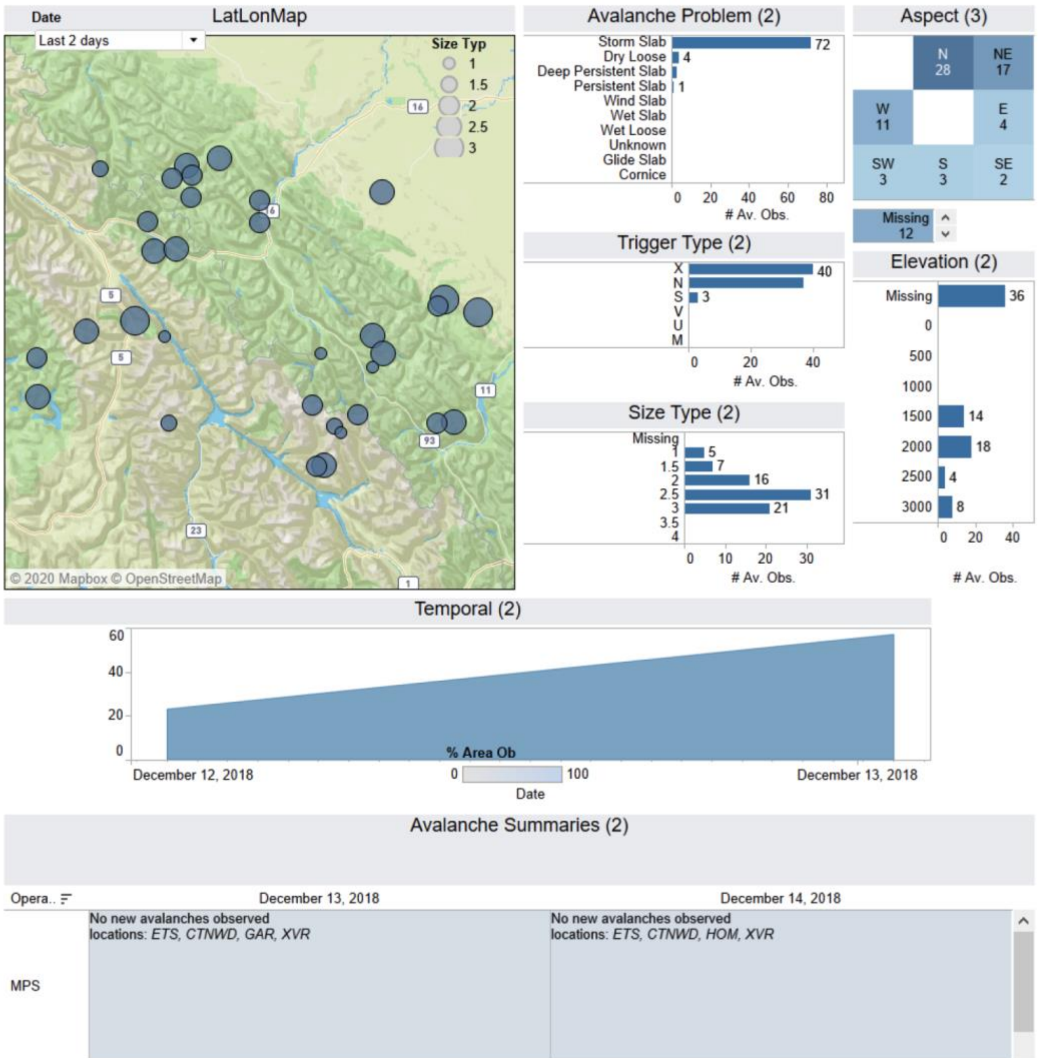


Figure 10 **An early prototype of the AvObs system utilizing precise visual encodings and numerical aggregates in the form of bar charts.**

Using the same conventional visualization design principles, our preliminary designs relied on numerical summaries such as counts and averages to provide an overview of avalanche activity (Fig. 10). We used simple bar charts to show summary counts of observed avalanches, but forecasters found this representation confusing, and they found it to be an impediment to their analysis. It did not support the demands of the task and forecasters found aggregates methodologically flawed.

> I like seeing the individual events more than the aggregate... It seems like full of flaws and limitations to kind of summarize all the [avalanche] activity with one number... (P11)

Moreover, forecasters expressed concern that very precise visualizations like bar charts would impart a **false sense of precision** in the data and have forecasters forfeit the scrutiny that such data require.

We learned of several reasons why forecasters find the disaggregated representations to be important. The first challenge stems from how avalanches are assessed. While assessing avalanche type and size using avalanche observation data is relatively easy as these are distinct data fields, judging the likelihood of avalanches is more difficult because it relies more heavily on expert judgment. Likelihood is formally defined as a combination of how sensitive instabilities in the snow are to triggering and the spatial distribution of the instability. However, as avalanche observations follow a targeted sampling approach, it is much more difficult to assess the representativeness of data. Spatial distribution, for instance, formally requires judging the "ease with which evidence is found" and specific terrain features or locations where the problem is expected to exist. This requires forecasters to assess the context within which observations were made, the areas traveled to, factors such as the mode in which observations were made, and the opportunities for observation. For instance, observing mountainous terrain on a clear day with good visibility while flying in a helicopter allows one to see much more terrain and avalanche debris than when flying on a cloudy day or if travelling on skis or a snowmobile. This contextual information can only be discerned by reading details of multiple attributes in each report. Moreover, more explicit and well-structured attributes of individual reports have to be viewed simultaneously to contextualize the meaning of any broader patterns in avalanche activity and thus how avalanche problems can be conceptualized. As a result, reports should be represented in a disaggregated form allowing forecasters to see multiple dimensions as they pertain to individual reports and allow more detailed information to be easily accessible.

The second challenge deals with ambiguous expressions of avalanche data where the application and interpretation of even structured data may be situationally contingent and potentially equivocal. While varying practices in how to classify certain avalanche problem types (Hordowick, 2022; Klassen, 2013) are one example of this. Another that emerged dealt with the number of observed avalanches field. This field may be reported using numbers corresponding to a count of observed avalanches or a qualitative ordinal field such as ‘several’ or ‘numerous’ that corresponds to specific ranges of how many avalanches were observed (Canadian Avalanche Association, 1995). However, the choice of which to was used and in which context leads to interpretations beyond a literal reading of the data. For instance, using the qualitative ordinal field could express uncertainty or simply be a time-saving shorthand. In addition, the number of observed avalanches has to be compared against the spatial scale in which observations were made. This is why forecasters were concerned about a **false sense of precision** or a face-value reading of data.

Glyph-based Design

We instead chose a glyph-based design to address the forecasters' concerns about false precision and trust (Fig. 11). Glyphs, composed of circle marks representing individual avalanche observation reports in a circle packing layout, are a central feature of our visualization design. Circle colour corresponds to the avalanche problem type. We chose this encoding because identifying avalanche problem types is generally thought of as the start of the hazard assessment process (Statham, Haegeli, et al., 2018) and colour is the most salient visual feature in glyphs (Borgo et al., 2013). Meanwhile, avalanche size is encoded as circle size. Avalanche type and size provide two essential elements for hazard assessment; to address the third - likelihood - we provide tooltip interactions (Fig. 11 & 12.I) showing many other detailed data attributes and unstructured data such as comments or images that allow forecasters to better understand the broader context within which to interpret observations. Finally, the data source is encoded as border stroke blurriness where recreationist reports (MIN) have a blurred border while professional reports (InfoEx) have a solid border.

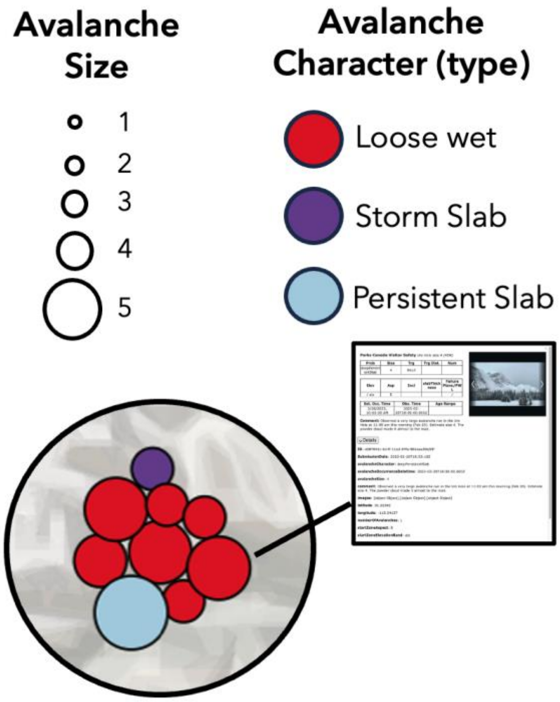


Figure 11 Glyph design using circle packing layout. Reports are shown as distinct circles allowing tooltip interactions that provide access to distinguishing information. Multiple encodings support visual aggregations.

Glyphs are embedded in a variety of visualization contexts. They are shown in a map (Fig. 12.B) where each avalanche observation may be associated with the tenure within which an operator travels (vector polygon) or specific locations, individual avalanche paths for instance, within the tenure (also vector polygons). Selections of circles show underlying spatial data color coded whether it is a tenure (green) or a location (yellow). Region boundaries for which target public forecast bulletins are issues are shown as contextual reference structure in purple. The circle packing algorithm prevents occlusion while zooming resolves reports to the associated spatial locations. They are also shown in a matrix grid (Fig. 12.C) organizing avalanches by avalanche type and various trigger types. This provides insight into subsets of different avalanches and their common triggers. In the design process, we invented new high-level trigger type categories departing from operational standards and better aligning with the analytic questions forecasters have. Namely, whether avalanches were triggered by natural causes, by humans in a deliberate controlled manner, by humans by accident, by explosive charges, or through some other means. They are also shown in a timeline (Fig. 12.E) that separates avalanche reports based on whether the avalanche was seen in motion (Direct Obs) or if only avalanche debris was observed (Age Range). For avalanches with an age range estimate, an estimated day in which the avalanche occurred is derived and the corresponding circle is aligned with that date. Selecting and highlighting observations reveals lines corresponding to the potential time interval in which avalanches could have happened. This helps forecasters understand temporal uncertainties and trends in avalanche activity.

A radial heatmap (Fig. 12.F), the only aggregate visualization showing numerical summaries, shows the number of reported avalanche observations by the aspect, the cardinal direction of a mountain face, and the elevation (in 200m bins) of where an avalanche was observed. This provides insight into the localization of an avalanche problem informing judgments of spatial distribution and in turn likelihood.

Finally, unstructured text fields called avalanche summaries that operators use to summarize avalanche conditions at a high level, are available in a scrolling list pullout (Fig. 12.D).

All views are linked and coordinated. Clicking axis labels, individual reports, or using the highlight lasso (Fig. 12.A) cross-highlights reports of interest in all displays. A lasso can also be used to filter data using the map and the colour encoding for glyphs can be changed between avalanche type (avalanche character) and the number of observed avalanches in each report. A separate pullout menu sets query parameters such as the date range, regions of interest, or avalanche problem types (Fig. 12.A).

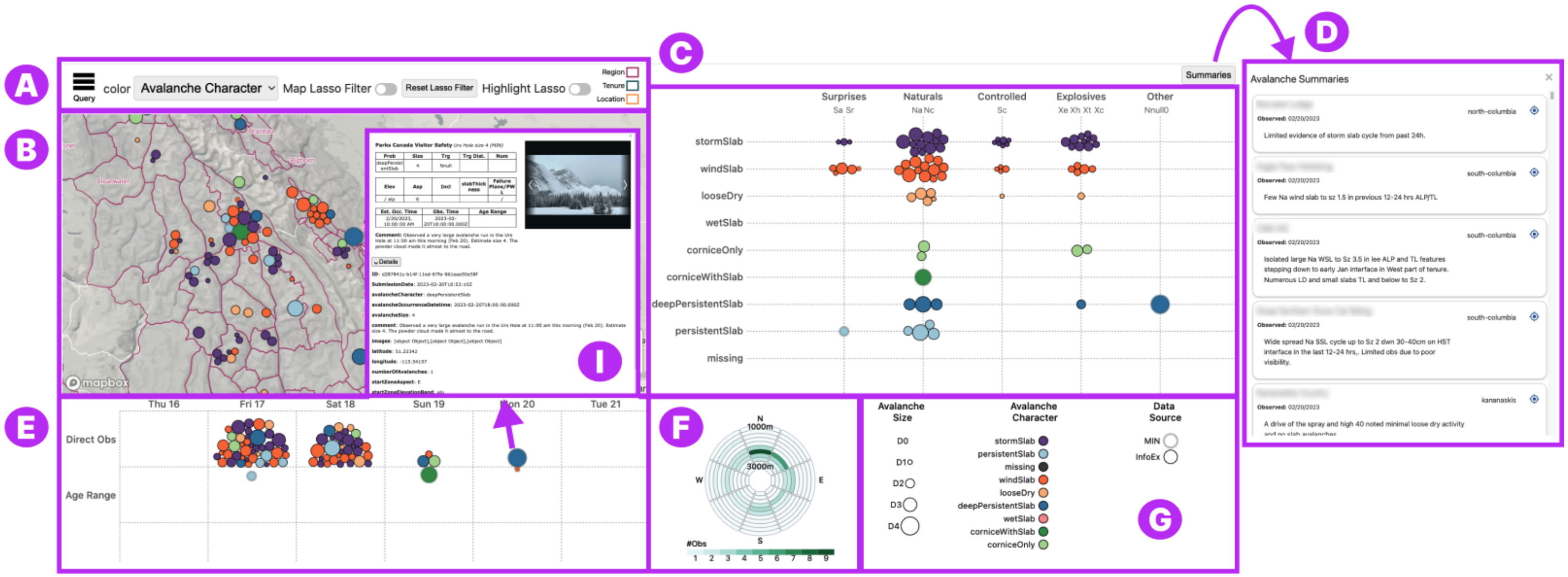


Figure 12 The AvObs prototype. A) Query parameters and global configuration such as a filter lasso. B) Map showing reports at their respective locations. C) A matrix showing subsets of reports by avalanche problem type and trigger type. D) Avalanche summaries are shown in a scolling list of text cards. E) A timeline showing reports by day and whether they were observed in motion or as debris. F) A radial heatmap showing avalanches by elevation and aspect. G) A legend for circle-packed glyphs. I) tooltip interactions to show report details.

Desirable Difficulty

We chose this glyph-based design to ensure forecasters had access to individual disaggregated reports while at the same time analyzing broader patterns and trends. Glyphs support visual aggregations such as summarizing data, detecting outliers, detecting trends, or segmenting data into clusters (Szafir et al., 2016). However, layering multiple encodings in a variety of contexts makes reading patterns in such visualizations more difficult (Healey & Enns, 2012). This was a deliberate design decision.

Forecasters expressed concerns about visualizations giving them a **false sense of precision** so we chose a visual design that would instead inspire scrutiny by making the reading of visualizations more difficult and effortful than our initial designs focusing on precision and salience. While size and colour are salient in glyphs, in the broader space of visual encodings they are not decoded as accurately (Cleveland & McGill, 1984) or perceived as saliently as position (Szafir et al., 2016).

The use of cognitive difficulties is often discussed in studies of learning (Chi, 2013; Yue et al., 2013) and is also present in visualization research (Hullman et al., 2011). It has been shown to improve risk-based decision-making in geovisualization applications (Cheong et al., 2020). Hullman frames "desirable difficulty" as involving a tradeoff between the cognitive efficiency derived from pre-attentive processing and improved learning through more active processing of information. By reducing the ease and fluency of decoding, more active attention is devoted to the task allowing inferences about missing information or inconsistencies that might otherwise be ignored. While desirable difficulty has commonly been applied to improved learning outcomes, we find it appropriate when dealing with ambiguity. Instead of relying on pre-attentive processing, where patterns in data may be taken at face value and processed in a more automatic fashion, difficulties require more effortful consideration and in doing so ensure that alternative interpretations are considered, and the data is appropriately scrutinized.

While our goal was to introduce some form of visual disfluency, forecasters felt comfortable with this representation as they felt it matched well to the analytic demands of their work.

> [The visualization] helps to smooth the data... and just at a glance... but it’s not smoothing where I can’t then... tease out nuances... I feel

> like it's really true to the data, which is a collection of individual points, kind of disparate points from across a forecasting region... (P2)

Forecasters also commented on how the visual metaphor of disorderliness matched the constraints of the data, especially when compared to prior designs (see appendix A).

> [The packed circle glyph design] really resonates… how it appears to put a lack of order to the observations. The previous one [referring to prototype using a dot plot] we saw [the data] seems way more structured than it actually is. [The glyphs] capture kind of randomness of that [data] (P11)

6.7. Discussion

Through our collaboration, I learned how important disaggregated views and access to raw details are for these data. Whether the data are sourced from automated remote weather stations or human field reports, forecasters need to see the details of individual reports to consider what factors are influencing the data and how to interpret them. Whether they are sensor errors or understanding the broader context of how data were sampled, these details are critical for interpreting the meaning and broader implications of these data. Forecasters must extrapolate from points across space and time, but to do so, they have to first understand the relationships between data and the phenomena or influence on avalanche hazard they represent. This relies heavily on knowledge, experience, judgment, and imagination.

Our design of AvObs breaks away from classic visualization conventions that value minimalism and precision. In doing so, it departs from the effectiveness principle. Precise, easy-to-read, and minimalist designs can impart a sense of authority or objectivity (Kennedy et al., 2016), and we felt this was inappropriate.

Rather than focus purely on abstracting data and mapping it to optimal visual variables, our design attempts to visually and metaphorically capture the imprecision and ambiguity of these data. One example of this is using a force-directed layout in the map view. It not only deals with the problem of occlusion when reports are spatially clustered but also alludes metaphorically to the fact that the precise location of the observation may not be known. At the same time, this level of precision might not be relevant at the broad spatial scale public forecasters work at. Zooming in on the map can resolve what the exact underlying spatial data is, but at lower-zoom levels, this spatial ambiguity is captured metaphorically through the jittering motion and repositioning that result from the constraints of the layout algorithm.

In addition, every time a selection is made, each glyph visualization is spatially re-organized in a random manner creating the impression of disorderliness and imprecision. The visual complexity and perceptual interactions between multiple layered visual variables could create different impressions of patterns between successive interaction states of the visualization. This unstable percept, which even in a static form, forces the viewer to strain to see the patterns, encourages scrutiny and skepticism about the patterns and relationships they believe to be perceiving. In this way, encoding is not just about mapping literal and explicit dimensions of data to visual features to extract summary statistics and patterns. It is also about aligning the analytic demands and affordances of a data set with a particular mindset or set of mental operations within a specific context. Rather than offload cognitive work and reduce effort, as is commonly thought of as the primary role and benefit of visualization, we are altering cognitive work. By making some aspects easier and automatic and others harder and more deliberate, our visualization aims to better meet the demands of the task at hand.

6.8. Limitations

Interviews with forecasters on the use of WxObs in real-world applications provide an ecologically valid perspective on the value it provides and its shortcomings. However, retrospective interviews are inherently limited due to issues of post-hoc rationalization, limited memory or understanding of the procedural aspects of cognitive work and how it relates to visualizations, and potential experimenter expectancy biases. In addition, while AvObs was designed through an iterative approach involving domain experts and continuous evaluation, a more rigorous and thorough evaluation of how design features do or do not address the challenges of ambiguity is needed. Moreover, this tool was not validated through operational use in this study leaving it in need of real-world validation.

Chapter 7. Study 3: Evaluating Visual Analytics in Practice

> *A map is not the territory it represents, but, if correct, it has a similar structure to the territory, which accounts for its usefulness.*
>
> *(Alfred Korzybski, Science and Sanity)*

I carried out a qualitative simulation study evaluating AvObs and the tools forecasters conventionally rely on: the InfoEx and MIN. Given that AvObs was designed in collaboration with expert users, this offered some indication that the new system would function as intended to provide a comprehensive multidimensional view of the data, showing trends and patterns without imparting a false sense of precision, and allowing forecasters to access details to better understand the broader context and implications of reports. However, a more thorough evaluation of how AvObs and existing tools work in practice was needed.

Evaluation of sensemaking visual analytics has traditionally focused on 'insight-based' methods (North, 2006; Saraiya et al., 2006). Most often, such evaluations involve simulations of real-world tasks in a lab-based setting. This methodology treats the goal of visualization as producing emergent 'insights' and discoveries which are categorized, counted, and used as a benchmark for comparing the performance of different visualizations in terms of the quality, type, and volume of insights they support in a given amount of time. In addition, it is heavily influenced by models of sensemaking developed for exploratory analysis. This methodology, its summative evaluation aims, and theoretical framing do not align with the goals and sensemaking context of avalanche forecasting nor the research objectives of this dissertation. In this research, I adopt the stance of Andrienko et al. (2018) treating the goal of visual analysis to be some assessment, prediction, or decision, supported by a process of iterative mental model calibration and development. The aim of this evaluation study is to understand how visualizations serve these goals, not to measure how *well* or how *much better*.

To the best of my knowledge, there are no existing sensemaking evaluation methods that treat the goal of visual analysis in such a way. Moreover, the operationalization of insights in evaluation studies often tends to focus on the information that is decoded from visualizations. Instead, as outlined in cognitive psychology on the 'representation effect' (Smith et al., 2006), my focus is on evaluating how visualizations come to aid the cognitive work demanded by the problem domain of avalanche forecasting. More specifically, the objective of this study was to evaluate *how* these tools support or impede ambiguity-related sensemaking and avalanche hazard assessment more generally. This is a formative evaluation and is not concerned with comparing the performance of AvObs against existing systems. Given the complex and contextual nature of ambiguous sensemaking such an approach would be an oversimplification as it is not clear what would be measured. Moreover, the specific analytic affordances of both tools as they pertain to ambiguous sensemaking are not known as they have not been studied in detail. As I could not rely on an existing methodological approach, I had to develop one to address this research objective.

7.1. Research Questions

1. How do AvObs and existing tools come to support or impede ambiguous sensemaking?
    1. What types of ambiguous sensemaking does each toolset support and how do they compare?
    2. How are features in AvObs and existing tools interacted with and how do they compare?
    3. What analytic support do AvObs and existing tools offer and how do they compare?
2. Do hazard assessments produced by forecasters using AvObs deviate from those produced using existing tools in a meaningful way?

7.2. Method

I carried out a qualitative simulation study in which we examined how forecasters used both their existing InfoEx toolset (Fig. 13) and the AvObs prototype (Fig. 12) in the same representative practice conditions.

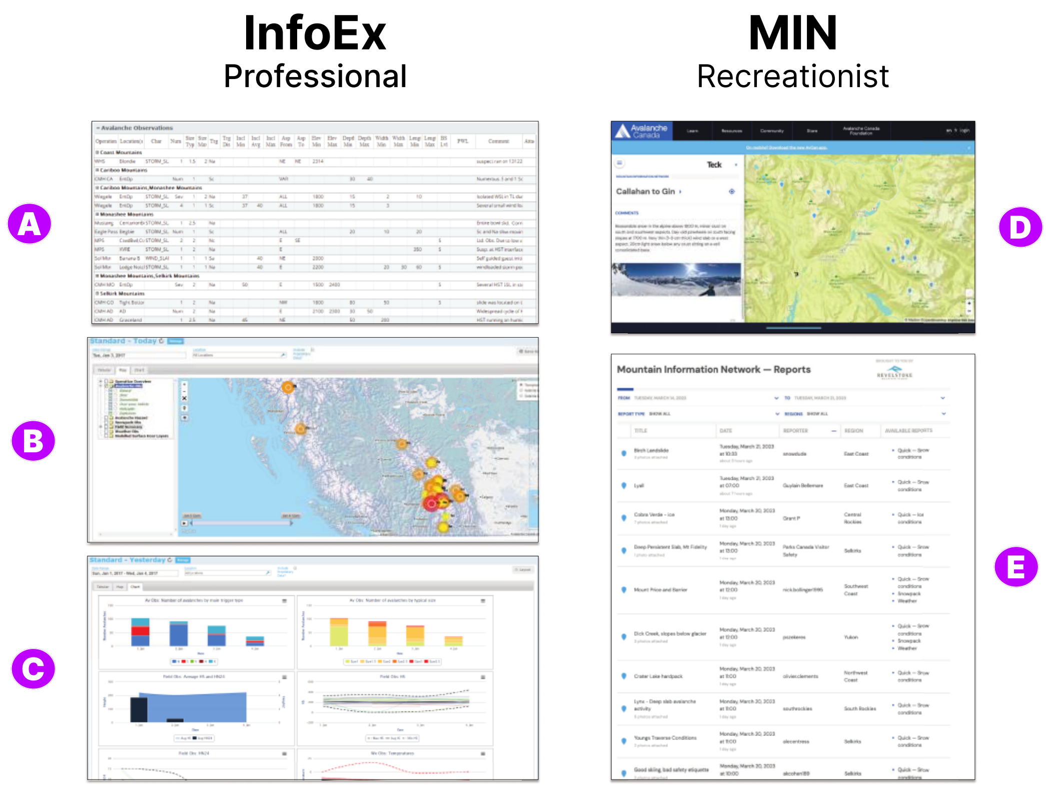


Figure 13 **Existing sytems used by forecasters in the analysis of observed avalanches. These include the professional InfoEx reporting system including unlinked A) data tables, B) maps, and C) information visualizations as well as the recreationist MIN reporting system including D) a map view with an interactive sidebar for accessing reports and E) a list view of reports.**

I asked our forecaster participants to conduct formal hazard assessments using either AvObs or InfoEx tools in two avalanche risk scenarios using historical data. Forecasters were asked to produce a 'nowcast', an assessment of conditions of the day in question, rather than a forecast because this relies more on human field observations and less on weather forecasts which would introduce additional complexity in our study. Hazard assessments included all formal measures that are conventionally used in North America including danger ratings, avalanche types, sizes, likelihoods (including sensitivity and spatial distribution), and a headline statement to address the public. Prior to the study, forecasters completed a guided walkthrough to gain familiarity with AvObs features. A think-aloud protocol was employed to capture vocalizations of thought processes and reasoning and to offer insight and explanation of actions taken using tools. The study was run remotely and online. The study was open to participants for one month, allowing forecasters to conduct the study at their own convenience and take as much time as they need.

7.2.1. Simulation Environment

In consultation with P2, who did not participate in this study, I designed two distinct scenarios constructed using historical data from prior seasons. I chose two Canadian forecast regions where there are relatively more professional operations and thus more data. This was to ensure that assessments were more based on the data featured in the study toolset rather than other factors. In addition, as avalanche conditions and the associated complexities and challenges of hazard assessment change over the course of the season, an early and a mid-to-late season day was chosen to capture a representative but varied set of typical forecasting situations. None of the participating forecasters had issued the forecast bulletin for the day and region I selected ensuring they would not be familiar. Even if they had seen the data previously, it is highly unlikely that they recalled or recognized it as both scenarios present very common situations and P2 checked the data for any reports that would be memorable. As the data used in these scenarios are proprietary, they cannot be published and disclosed. However, Figure 12, which shows an image of the AvObs prototype, provides a representative view of data from a typical forecasting scenario.

In keeping with resources forecasters conventionally have available, I provided forecasters with the bulletin and forecast issued on the day previous to each scenario. Participants were also given synthesized weather data from weather stations for current conditions in the scenario region.

7.2.2. Conditions

I divided forecasters into two groups, such that each scenario (S1 & S2) was used in both InfoEx (using both MIN and InfoEx systems) and AvObs (integrating MIN and InfoEx data in a single system) conditions to control for any difference arising from the scenario rather than the tool. While presentation order was not a major concern, InfoEx was always presented first in case AvObs inspired any alternative uses of InfoEx that deviated from those conventionally used:

- Group A: InfoEx + S1 → AvObs + S2
- Group B: InfoEx + S2 → AvObs + S1

7.2.3. Participants

10 forecasters participated in the study (Table 7). Six were forecasters at Avalanche Canada and four were forecasters at various other organizations throughout North America. They varied in experience from 1-24 or more years and professional backgrounds as engineers, scientists, mountain guides, and educators.

Table 7 Participant table for Study 3

ID	Public Avalanche Forecasting Experience	Background	Public Forecasting Organization
3	4+	Geological Engineer	Avalanche Canada
4	10+	Mountain Guiding, Ski Patrol	Avalanche Canada
5	5+	Mountain Guide, Educator	Avalanche Canada
6	19+	Mountain Guiding	Avalanche Canada
7	6+	Ski Patrol, Mountain Guiding	Avalanche Canada
13	1+	Avalanche Field Technician	Avalanche Canada
14	11+	Mountain Guiding, Engineering	Colorado Avalanche Information Center
15	12+	Mountain Guiding, Geologist	Parks Canada
16	19+	Geographer	Colorado Avalanche Information Center
17	5+	Engineer	Avalanche Canada / DAC

7.2.4. Data Capture

Participants video-recorded their screens and vocalizations of their thoughts while completing the task. Additional study materials were disseminated remotely using SurveyMonkey[9]. Materials were presented in sequence. First, forecasters were presented with a participant consent form and a questionnaire asking about their professional background, years of experience as a public forecaster and as a forecaster in other contexts, and the public forecasting organization where they work. Next, each study condition was presented in order. Each condition provided instructions on how to

[9] https://www.surveymonkey.com/

record their screen, vocalize thoughts and explain actions, which tools to use and the span of allowed data to query, and general context for the task and scenario. In addition, the prior day's bulletin and current weather data were made available followed by questions in which to enter formal hazard assessments. Finally, a questionnaire soliciting feedback about the AvObs tool was presented. Questions asked about features forecasters found most and least useful, any functionality that is present in InfoEx and missing in AvObs, and specific feedback about how spatial, temporal, and informational aspects of the data were represented (More details in Appendix A.)

7.2.5. Analysis

We conducted video analysis by thematically coding videos for sensemaking processes, interaction logs, and analytic actions. As will be expanded on in detail below. a mix of inductive open coding, theory-driven deductive coding, and coding involving provisional starting codes was used at various stages.

Sensemaking Processes

We coded sensemaking according to two features: the process and the context. We adapted a coding structure using sensemaking processes from the data-frame sensemaking theory (Klein et al., 2007) and operationalized ambiguity-related sensemaking as instances where frames were either questioned, compared, or reframed (distinguishing it from conventional sensemaking involving mapping frames and data as well as augmenting existing frames). Frames included any explanations or expectations about conditions, data, or any other topic matter the forecaster discussed. For instance, explanations for assessments made in the prior bulletin or expectations about the shape and characteristics of data given the forecasters' current understanding. This was then used to identify distinct types of ambiguous sensemaking which emerged from study data, but also aligned with findings from Study 1.

In addition, the context within which these sensemaking processes were applied was captured. These fell into two types: *hazard analysis*, in which forecasters are analyzing the observations and data, and *hazard assessment*, in which they are refining their assessments in *the bulletin* by integrating information from the prior day's bulletin along with their own hazard analysis. The context was determined not only based on which window the forecaster was viewing but the content of what they were talking about. For instance, if they were planning how to characterize assessments while viewing an analytic tool, this was considered to be hazard assessment and not hazard analysis.

The codes pertaining to data-frame sensemaking processes and contexts were applied and refined in two passes by one coder. Lower-level ambiguity-related sensemaking processes were developed in one pass and refined in two passes by one coder. Reliability for all codes was assessed concurrently by two coders applying the code structure to three videos. Areas of disagreement were discussed and the coding guide was adjusted to alignment. Codes were applied using MAXQDA[10].

Interaction and Analytic Action Codes

To capture how tools were used in analysis, we coded for low-level interactions. As proprietary tools used by avalanche forecasters could not be instrumented and much of how these tools are used relies relatively more on visual information than mechanical actions, we treat interactions as any form of tool-use, including the manner in which visual information is viewed, and focus on information that can be inferred in video recordings from participant vocalizations and visible behaviours, such as cursor movement. We were interested not only in what was clicked but also which views and charts were used and how. This included the tool being used; the chart or view being used; explicit interactions such as filters, selections, or zooming and panning; information-seeking behaviours such as scanning across or down columns; or the investigation of trends and patterns in available charts. Three coders each independently developed a coding structure for one video in two passes. Two coders were each allotted five of ten participant videos for coding. Reliability was assessed by comparing codes for four videos, two from each allotment. Areas of disagreement were used to adjust coding guidance. Codes were applied using Elan[11], a commonly used video analysis tool.

I then used interaction codes to infer higher-level generic *analytic actions*, drawing from research on analytic tasks in visualization (Brehmer & Munzner, 2013; Rind et al., 2016; Shneiderman, 2003) and cognitive theories of reading (Elfenbein,

[10] https://www.maxqda.com/

[11] https://archive.mpi.nl/tla/elan

2018). Analytic actions are defined as domain-agnostic steps using visualization to address analytic objectives (Rind et al., 2016). My intent was to use these abstractions to enable comparisons, if only qualitatively, about how tools might support the forecaster's analytic needs. This yielded five analytic actions (described in Table 8) from observed behaviours, vocalizations, and interaction codes.

Table 8 Observed analytic actions, their definition, and the evidence used in identification.

Action	Definition	InfoEx Interactions	AvObs Interactions
Drill-down	Identifying data items of interest using visual information about data attributes to access details	Grouping data by attributes using table sorting and mouse cursor to identify subsets of reports of interest and access details by reading.	Using chart layout, visual encodings, or cross-highlighting to identify reports of interest and access through tooltip interactions
Serial	Identifying data items of interest purely through presentation order or sequence and then accessing details.	Reading each row in the table in the order it is presented in.	Opening tooltips for reports in the exact order they are presented in or reading each text block in the conditions text synopsis list in sequence.
Pattern	Investigating high-level patterns like trends, distributions, or proportions.	Scanning down table columns using mouse cursor and talking about distributions, trends, minimums, maximums, or central tendencies.	Mousing over charts and talking about trends, distributions, spatial distributions etc.
Focus	Spending a disproportionate amount of time investigating a single data item of interest.	Spending a disproportionate amount of time reading or re-reading a single table row.	Keeping tooltip open for an extended period or expanding tooltip for further details.
Gisting	Rapid and automatic processes for apprehending the core meaning from complex visual information.	Cursor jumping across table with no discernible pattern and no indication of how information search is being conducted. Participants often murmuring under breath.	No observed examples.

One researcher applied the codes to all videos in two passes using video playback for context and to ensure the abstracted interpretation presented by an analytic action was coherent and made sense independent of codes captured in interaction logs. Codes were applied using MAXQDA.

Result Metrics

I used two dependent measures to assess the results. First, the formal hazard assessments forecasters produced served as a measure of analytic outcome. This could provide an indication if AvObs biased forecasters' judgment in any way. Given the small sample size, and the considerable variability of forecaster's hazard assessments under normal conditions (e.g., Statham, Holeczi, et al., 2018), I looked for meaningfully large differences with plausible explanations.

Second, simple descriptive measures were used to summarize the thematic codes (e.g., observed actions): median counts of a thematic code showed the typical number of times a participant demonstrated a particular behaviour, the number of participants exemplifying the coded behaviour provided an indication of how common behaviours were across the sample, and the average duration of each coded activity gave an indication of the relative time spent on different activities. Together, these simple measures provide some insight into the relative effort devoted to certain activities and thus can serve as an indication of the relative importance of different tasks, information, and tool features and the propensity for either toolset to support different ambiguous sensemaking tasks and analytic actions. Similarly, as these measures represent a qualitative coding approach and are thus subject to errors, only meaningfully large differences with plausible explanations were considered.

7.3. Findings

When comparing toolsets, I did not find meaningful differences in the hazard assessments forecasters produced (Fig. 14). All measures of hazard assessments were similarly variable and similarly distributed in both scenarios for both tools. However, I did find differences in how often forecasters engaged in the different sensemaking activities and analytic actions.

Figure 14 A comparison of hazard assessments between conditions in both scenarios does not show meaningful differences. A) The distribution of danger ratings according to vegetative bands *(BTL= below treeline, TL = at treeline, ALP = Alpine).* B) The order and perceived importance of avalanche problems by participant. C) The primary avalanche problem for each scenario showing likelihood and size assessments. This includes a minimum and maximum value (bounds of coloured bars) and a typical value (black circle). D) The secondary avalanche problem for each scenario.

7.3.1. Ambiguous Sensemaking

I observed a variety of ambiguous sensemaking processes involved in both hazard analysis and assessment. A key distinction arises between the AvObs and InfoEx toolsets during hazard analysis (Fig. 15). A larger number of forecasters engaged in ambiguous sensemaking processes when using InfoEx, and ambiguous sensemaking was also overall more common with the InfoEx toolset. However, the relative proportions of types of processes appear consistent. Some sensemaking processes appear to carry more significance for tasks related to hazard analysis, while others are more related to the needs of hazard assessment.

Forecasters **speculated** about factors explaining data information. For instance, in hazard analysis, some conjectured about how physical processes or data sampling procedures influence how data is shaped.

> most [avalanche activity] that's been reported is treeline and below but that also might be because that's just where people are skiing... it's storming they can't land in the Alpine. (P10)

During hazard assessment, speculation often had to do with understanding fellow forecasters' reasoning processes.

During analysis, forecasters frequently **mentally converted** data values to account for biases, inconsistencies, or extrapolations across space and time. A common example involved questioning the classification of observed avalanches after reading the details of the report.

> Depth 40… that tends to make me think that was actually a storm slab problem [as opposed to a different avalanche problem type] (P17)

In another example, P6 used their prior knowledge of a locale to adjust a bias in size estimates.

> a 1.5 [avalanche size] in the Monashees [addressing operators in a particular area] is at least a two (P6)

Mental conversion was considerably less common in hazard assessment and dealt primarily with interpreting formal assessment standards and the prior bulletin.

I observed how forecasters **weighed evidence** during analysis based on diagnostic strength and representativeness. P10 used his knowledge of localized terrain to discount a set of large, reported avalanches because they were not representative of broader avalanche conditions.

> triggered two and a half's from Gullies on Mt. McDonald, I pretty much discard those avalanches on that huge mountain because they're always two and a half's and threes (P10)

Meanwhile, P17 put more weight into a report authored by a recreationist they were familiar with and trusted.

> He's one of the power users in Rogers pass and I actually put more weight on his observations than I would on the average MIN user (P17)

Weighing evidence was considerably less common in hazard assessment.

Finally, **triangulation** - a process of refining understanding through comparison with alternative perspectives - was common in assessment but not analysis. To better understand how to assess conditions, forecasters referred to the information provided in the prior day's bulletin, formal documentation and guidelines to compare with their subjective perceptions, and their own prior assessments check for soundness, coherence, and self-consistency. Rather than simply averaging multiple judgments for a more accurate prediction, they used this to develop coherence and identify new issues to consider.

> Assuming we are talking about the new smaller wind slabs, I don't know if I agree with the step down part here… unless it's the wind slabs that are directly sitting on this, the step down is covered by the persistent problem not the wind slab problem (P17).

The path from hazard analysis to assessment is not sequential; instead, forecasters constantly iterate between the two contexts (Fig. 16). While these processes are clearly closely coordinated, my focus is primarily on evaluating how analytic tools support the demands of this work. To this end, I now shift attention to processes important in hazard analysis and how tools may come to support the processes of mental conversion, weighing evidence, and speculation. In the following section, I examine the interactions and analytic actions each tool affords to inform how they might support ambiguous sensemaking processes.

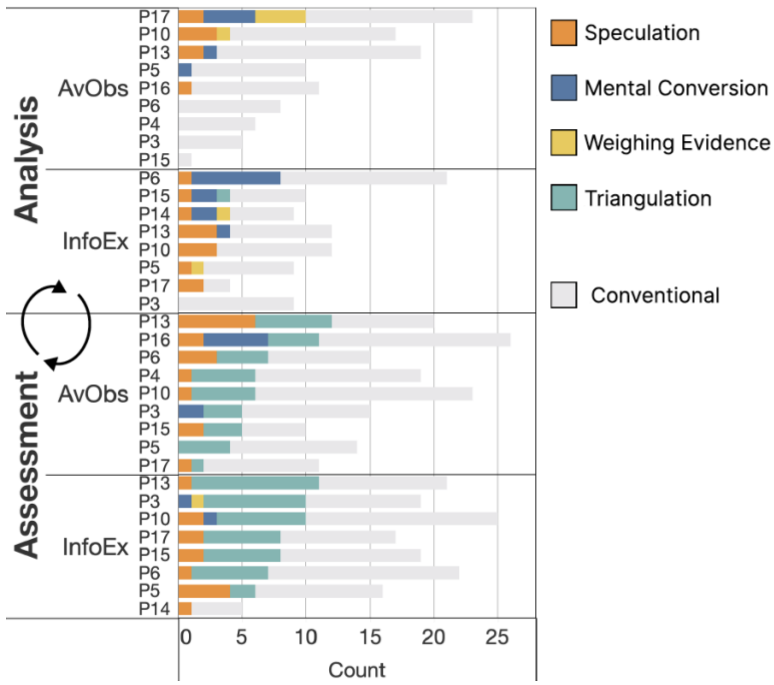


Figure 15 **Counts of ambiguous sensemaking tasks by sensemaking context.**

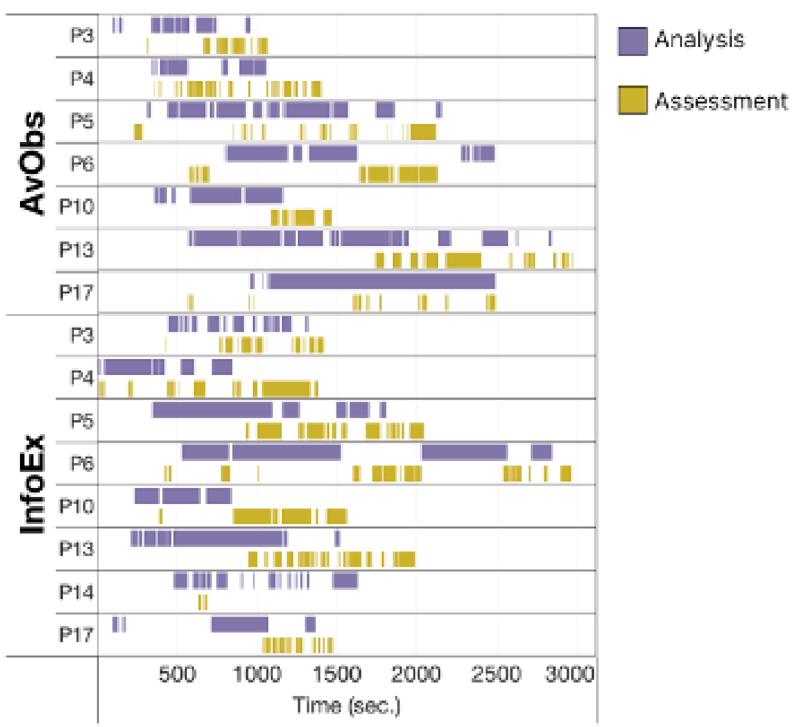


Figure 16 **Sequences of activity are shown by sensemaking context.**

7.3.2. Analytic Actions and Tool Affordances

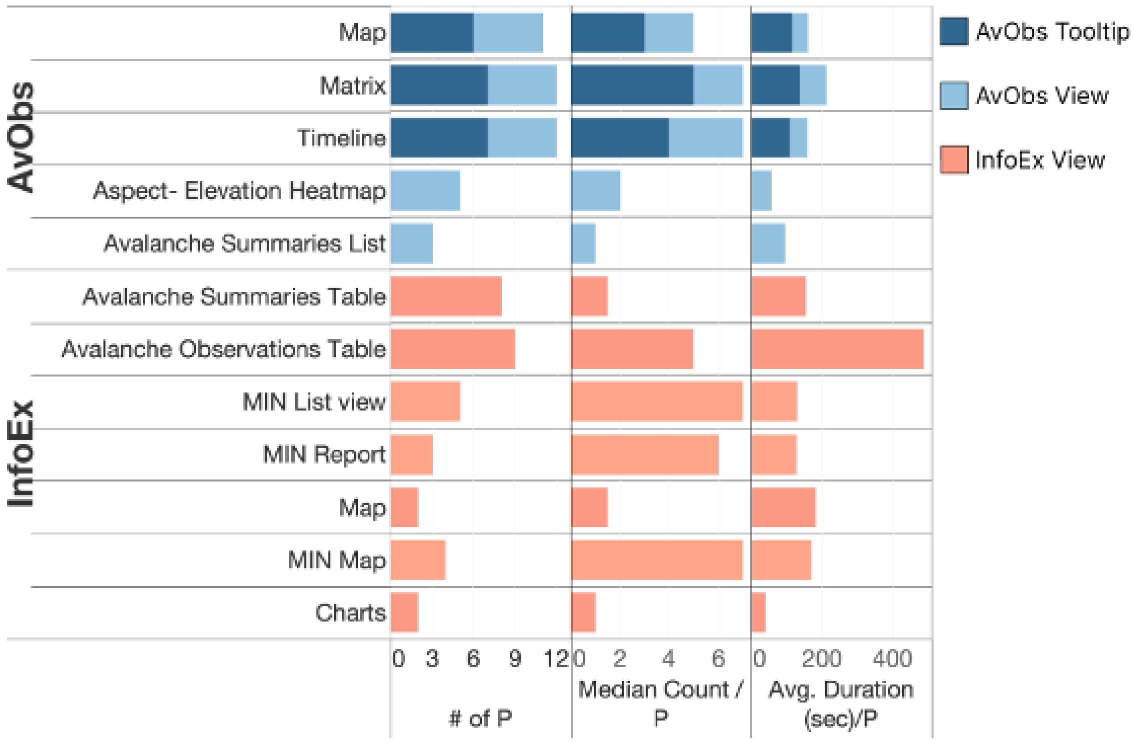


Figure 17 **Participants' use of various views in the InfoEx condition (red) and the AvObs condition (light blue). Interactions involving reading patterns and tooltips (dark blue) are shown for core AvObs visualizations illustrating the relative effort devoted to accessing detailed information.**

I found differences between the tools in how much effort was involved in finding information of interest and how representations support navigation to this information. Specifically, forecasters expended much more effort toward detailed information within reports rather than high-level patterns across reports. In addition, differences in interactions and analytic actions reveal differences in the analytic affordances of either set of tools.

This is apparent in how forecasters used different views. Charts and maps were used considerably less frequently than text table views in the InfoEx system and the InfoEx condition overall (Fig. 17). However, maps in the MIN system (part of the InfoEx condition), were used more frequently by forecasters who used the MIN system. However, the MIN system does not feature a comprehensive view showing all attributes of reports simultaneously, as in the InfoEx tables, and the MIN map and list views were used exclusively for accessing the details of these reports rather than investigating broader spatial or aggregate patterns.

In AvObs, I observed forecasters using available charts quite frequently. However, as highlighted in Figure 17, the use of tooltips to access detailed information which normally would be available in InfoEx tables was more common than investigations of high-level patterns. This reveals the relatively higher effort forecasters dedicated to investigating details of reports as well as the affordances of either toolset in navigating to information of interest. I explore these differences further through high-level analytic actions derived from interaction logs.

All forecasters *drilled-down* to details by first identifying reports of interest using visual information in the display (Fig. 18). However, forecasters using AvObs did this more frequently and for longer durations. In InfoEx, this involved sorting tables and subsetting reports of interest with certain characteristics and then accessing details by reading the full report row. In AvObs, this involved using the layout, visual encodings, and cross-highlighting interactions to identify reports of interest and then opening tooltips to read the report.

By contrast, reading reports *serially* in the order that the reports are presented without first identifying which reports are of interest was considerably more common in InfoEx. In AvObs, only three forecasters engaged in this activity, and only when reading free-form text avalanche summaries (Fig. 12.D) in the side panel, rather than opening tooltips for reports in the primary AvObs display. Forecasters using the InfoEx spent more time reading reports serially.

Overall, more than half of the forecasters were observed to be investigating *patterns* in either tool, but I found that investigation of patterns in AvObs seemed to prompt ambiguous sensemaking more often. One notable example using the single aggregate visualization in AvObs, the aspect-elevation heatmap, prompted speculation.

> There aren't really any persistent observations on direct south-facing slopes. So maybe the surface hoar layer got destroyed there? I don't know, just speculating… (P13).

InfoEx presents elevation and aspect in separate attributes making the task of synthesizing such patterns considerably more difficult.

Similarly, more than half of the participants in both toolsets spent time *focusing* on individual reports. In AvObs, where accessing some details involved not only opening a tooltip but also expanding a scrolling list of further attributes, I observed focus marginally more frequently.

Finally, the starkest difference was in *gisting*, the activity of getting "the big picture" of the information space. I observed this only when forecasters were using the InfoEx tables. I inferred this by a combination of behavioural cues: the forecaster's mouse cursor jumping around InfoEx tables, sporadically fixating on seemingly random cells, with no discernable visual or auditory indication of how forecasters were navigating the information space, and murmuring as they did this. While I was attuned to seeing similar behaviour in AvObs, there was no evidence of this activity.

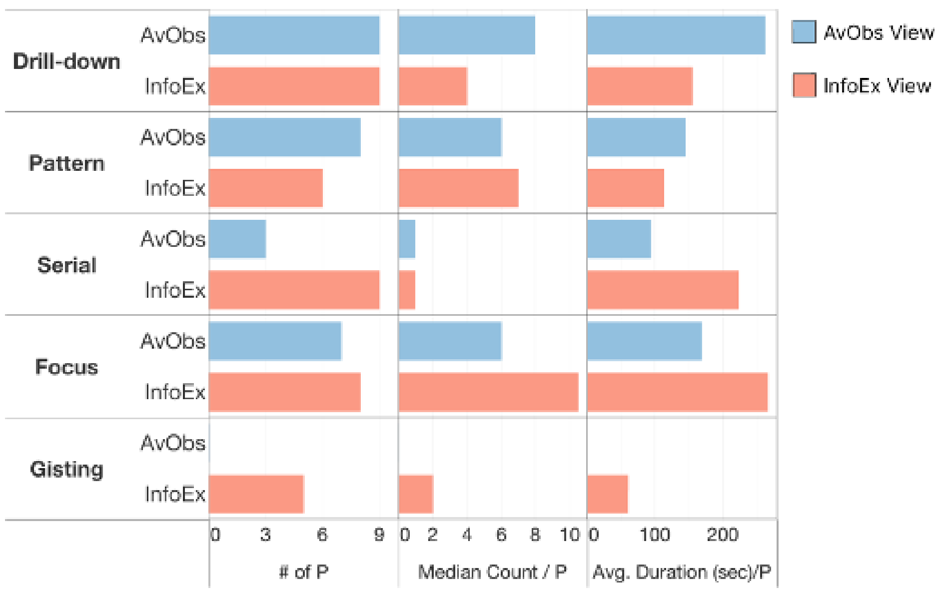


Figure 18 Chart comparing InfoEx and AvObs conditions according to indicators of effort devoted to various identified analytic actions.

7.3.3. Participant Feedback

In the questionnaire feedback, forecasters were generally enthusiastic about AvObs and its potential but expressed some concerns that AvObs was missing.

> The raw layout of data where a forecaster can make up their own thought on the distributions (P3)

They were also concerned visualizations may be misleading.

> perceive patterns in data when the sample size is too small (P5)

P5 also emphasizes the use of language in reports as a key indicator.

> How a day of avalanche activity is talked about is very important to me and often says more than the 'hard' data alone (P5)

P3 stressed that the textual avalanche summaries, which were used considerably less in AvObs (Fig. 13), were de-emphasized in our design.

> The summaries are important for seeing areas that didn't have activity or for other notable events not entered in the tables, but it is easy to forget the small box in the top corner. (P3)

Such information is critical for assessing likelihood.

7.4. Discussion

In summary, the key findings from this study are that either toolset did not bias the judgment of forecasters in a way that meaningfully affected hazard assessments, that forecasters engaged in more ambiguous sensemaking when using existing tools in the InfoEx condition, that forecasters dedicated a disproportionate amount of effort accessing detailed information, and that each toolset presented different affordances in terms of analytic actions. Namely, that InfoEx supported gisting while this analytic action was not observed in AvObs.

In retrospect, it is unsurprising that the hazard assessments produced in either toolset condition were not noticeably different – this the forecasters' job after all. I was encouraged that the AvObs prototype did not bias their judgment and lead to different hazard assessment outcomes. At the same time, findings in this study do offer some insight into how visual and interactive idioms can serve to support or impede sensemaking, and they illustrate that conventional approaches to visualization design have shortcomings in applications like avalanche forecasting.

InfoEx, largely constructed from the base textual data, seemed overall to have a higher propensity for prompting ambiguous sensemaking, whereas AvObs provoked concerns about access to all the data needed. Given that InfoEx is their current operating tool, one might expect simply that they are more accustomed to it. But I conjecture that differences in analytic affordances relate to assessing the nature of what is interesting, seeing enough detail in the larger context, and actually getting to the important information, highlighted strengths and weaknesses between the two approaches. In short, the differences in ambiguous sensemaking may have to do with differences in the accessibility of and emphasis placed on detailed information within either toolset.

I knew details (the raw "base data"), were important in this work, but I did not anticipate the scale of effort that is devoted to accessing this information. AvObs placed an additional interaction step towards accessing details whereas InfoEx tables make details available at a glance. While this may have hampered some forms of inquiry, AvObs also presented some advantages in supporting ambiguous sensemaking. It supported fluid integration between spatial, temporal, and categorical attributes of data and scaffolded insights about patterns otherwise difficult to understand. Findings showed forecasters using aggregate summary views to speculate about alternative factors that could explain the shape of data. Taken as a whole, the relative importance placed on detailed text which was only accessible through laborious and fragmented interactions is problematic for any task that relies on integrating information across tooltips. Any information accessed through tooltip interactions, bearing any relevance to the task at hand or relationship to details found in other reports, is transient, has to be remembered, and is not easily assembled with other such details during the flow of analysis. I speculate that this fragmentation and the relative importance of detailed information in this problem area may have contributed to why ambiguous sensemaking was more commonly observed in the InfoEx condition.

This may be because the gisting afforded by data tables in the InfoEx system makes apprehending the relevant information more fluid. Details are available at a glance and the structural layout of the table makes it easier to retrieve this information. Further, I saw how the layout of data tables allows other analytic actions such as drilling-down using attributes of interest or discerning patterns such as trends, distributions, and central tendencies by scanning down columns. I conjecture one of the main benefits of this gisting is how rapidly and fluently forecasters can switch between these actions using learned visual scanning procedures. The design of data tables does not take advantage of perceptual salience and attention-directing mechanisms in the same way that common visual idioms do. The visual hierarchy is much flatter. Consequently, the data table makes a much wider set of potential inferences from data possible and allows the observer to themselves direct attention where and when it is needed within the flow of analysis. Where data are ambiguous and comprehensive understanding of contextual details is important, this can be advantageous.

While the scale and volume of data in this application area make it possible to read all reports in a realistic scenario, this does not preclude the relevance of gisting in applications with higher volumes of data where reviewing all details is not possible. This analytic action has also been observed in studies of data workers and is applied in the interrogation of data regardless of scale (Bartram et al., 2021). Furthermore, the relevance of gisting in the context of ambiguous sensemaking has more to do with the accessibility of **relevant and raw** information from which potential explanations can be recognized, compared, constructed or refined. In cases such as avalanche forecasting, where data are ambiguous in that there are multiple potential factors explaining the shape of data and subsequent interpretations bear risk implications, gisting as an analytic action is especially important because the cost of a false negative is much higher than a false positive. In other words, ‘missing’ critical details that are difficult to anticipate can have serious consequences. While visualizations have the advantage of making it easier to see certain patterns and outliers that fall outside a certain way of framing data, they can occlude important details and in doing so, leave the viewer comfortably unaware.

I did not find that either AvObs or data tables in InfoEx are overall a better analytic tool for avalanche forecasting: each brings representational advantages. Since this evaluation study was run, AvObs has been incorporated into operational workflows at Avalanche Canada and several other forecasting organizations operating in Canada. From early reports and anecdotal evidence, I have learned that forecasters use AvObs to provide them with a quick summary overview at the beginning of the day, to review the prior day's forecast, and to have a first impression of broader patterns in the data. They then continue to use the InfoEx data tables for the subsequent detailed analyses throughout the day, to ensure that they are not "missing anything" in the text reports. A further investigation of how these tools are used in real-world applications is warranted. In addition, future work should further validate and re-evaluate the formative understanding that has been developed in the present exploratory study.

7.4.1. Re-evaluating Visualization Design Guidelines

I argue that conventional design guidelines are misaligned with the cognitive work involved in ambiguous sensemaking. The overviews and aggregations produced through strict adherence to the information-seeking mantra (Shneiderman, 2003) or precision-based visual variable *effectiveness* rankings (Bertini et al., 2020) can be misleading or at least inappropriate when dealing with ambiguity in risk-based applications. Visualizations can occlude important information (McNutt et al., 2020) and in doing so impart a misplaced sense of confidence in inferences made. As the costs of missing critical information are high and as evidenced by the concerns of forecasters in this and prior studies, this raises issues of trust.

Visualizations need to provide access to information of relevance. When nuances of how data are shaped have bearing on interpretation, as is often the case in analysis but especially is especially pertinent in risk-based applications, access to details in the raw base data is critical. The disproportionate amount of time forecasters spent navigating to and reading this detailed information is evidence of this. Moreover, access needs to fluidly integrate into the flow of analysis as meaning is being constructed. Simply providing access to details-on-demand through interactions is not enough because it could serve to interrupt this flow. The fact that the InfoEx resulted in a greater amount of ambiguous sensemaking serves as an indicator of this. In addition, anecdotal reports of the role AvObs and InfoEx play in real-world applications also support this understanding. Specifically, that InfoEx is used for core analyses ensuring forecasters are not missing critical relevant information.

I argue that visualizations need to better support *gisting*. A data table is a visual representation that clearly supports this distinct analytic action, but I don't believe it is necessarily the only one. I believe that the capacity to show the relevant data in its raw form and provide access to details to allow various processes like *gisting* could be supported by a variety of representations. Further, I don't believe that tables and visualizations are an “either-or”. While combining the two is not uncommon, I suggest this surfaces a larger challenge to the visualization community: how to better integrate visualization strengths of enhanced pattern and feature recognition with sequential reading and scanning strategies common to symbolic representations.

7.5. Limitations

As with all ethnographically-inspired research, this study is prone to limitations in reliability and validity. This study was essentially exploratory, and I acknowledge a number of methodological limitations. Evaluating what people are thinking in a complex reasoning task is inherently problematic and imprecise. In an ideal world, I would have liked to trace the relationship between specific interactions, analytic actions, and how these lead to specific various ambiguous sensemaking processes. However, drawing a one-to-one correspondence between these levels of abstraction is not possible in such a complex environment which also relies heavily on prior knowledge. Much of the information used in the forecaster's sensemaking is not explicit in data nor is it vocalized by forecasters making such an analysis intractable. Thus, my observations must be treated as indications for further questions, rather than conclusive proof of cause and effect.

I also note the shortcomings of using observations, screen recordings and forecaster vocalizations to capture interactions and attention. They only provide partial insight into how visual information is being used in analysis. Future work may use other methods such as eye-tracking to make more refined inferences about patterns of how visual information is being scanned. This may, for instance, provide more reliable information about analytic actions such as gisting.

Chapter 8. Study 4: Exploring Knowledge Capture

Everything simple is false. Everything complex is unusable.

(Paul Valéry)

While Study 2 & 3 address issues of individual sensemaking, they do not deal with the challenges of shared understanding identified in Study 1. In this chapter, I describe a set of preliminary investigations for processes of shared knowledge construction and how they can be supported to better facilitate collaboration and communication in this domain.

Using participatory design and diary study methods, the forecasters and I explore how visual analytics systems can enhance collaborative sensemaking in asynchronous hand-offs. Specifically, how domain knowledge may be used as a templated structure to simplify the process of gathering, organizing, and communicating materials for hand-off.

It is important to note that this study falls into the early stages of the iterative design and rapid prototyping phase of a visualization design study. It serves to develop a preliminary formative understanding by identifying core tasks, needs, challenges, and preliminary indications of potential solutions.

8.1. Motivation

In study 1 (Chapter 5) I identified asynchronous sequential “hand-off” of work between forecasters at shift changes as a critical challenge. As avalanche forecasting relies on continuity of analysis – with historical understanding of avalanche conditions having direct bearing on understanding the current and future states of avalanche conditions – hand-off is a challenge because communications tend to be incomplete and disruptive to a given forecaster’s more immediate daily goals and responsibilities. Forecasters use the prior day’s bulletin as the starting point for their workday. It is used to plan work, set expectations, and guide information-search. They iteratively update the bulletin, changing assessments as they work to bring them in line with current understanding. However, as the relationship between assessments in the bulletin and reasoning processes of forecasters is not captured, collaborators often have to try and reconstruct their collaborators reasoning processes by investigating the available evidence that could serve as an explanation.

In addition to the bulletin itself, forecasters at Avalanche Canada also employ a variety of other tools to facilitate asynchronous sequential collaboration. They use chat tools for direct communication, which help maintain a record of discussions about conditions, evidence references, and reasoning processes. Furthermore, these tools are also used for communications with field teams. Internally, forecasters utilize specific documentation to track weak layers in the snow and identify avalanche issues not mentioned in the public bulletin. Additionally, daily synchronous discussions with fellow forecasters also help fill gaps in understanding.

These materials are often compiled at the end of the day, as a separate and additional step beyond normal responsibilities. Consequently, they are often incomplete. This is problematic because materials used for monitoring evolving conditions should align with the state of current conditions. Further, this can lead forecasters who are starting their shift to misallocate their time. In Study 1, forecasters described a latency in how long it takes them to “come up to speed” and gain enough confidence to start making challenging assessment decisions.

I frame this as a challenge of distilling and articulating tacit knowledge, embedded within a particular working context. I draw inspiration from common ways to capture knowledge and document reasoning in other domains. One of the most common examples is the use of marginalia, or notes in the margins of books, to facilitate knowledge construction and critical thinking (Jackson, 2001). This goes beyond the concept of annotation and speaks to a variety of actions like highlighting or the development of common categories of information that facilitate thinking and the progressive development of emergent abstractions. This is a common practice in reading (Brummett, 2018), in taking notes, and is an interaction paradigm that is ubiquitous in interactive spreadsheet programs (Bartram et al., 2021). It is a practice whereby the observer begins the task of narrating the understanding they are developing as they read or analyze data, while still maintaining a trace of the originating context. Drawing inspiration from these other media, I explore what flexible knowledge support tools for avalanche forecasters might look like within a visual analytics tool ecosystem.

8.2. Research Questions and Objectives

I frame the challenges faced by forecasters as having to do with issues of shared context, coordination of work, and the principle of least collaborative effort as defined in collaborative visual analytics (Heer & Agrawala, 2008). **Domain knowledge can provide structure to better facilitate the capture and communication of analysis** (Rind et al., 2019) **offering a schematization mechanism for meta-data.** We sought to characterize these knowledge structures and investigate how they could be implemented in a visual analytics system for both the capture and communication of analysis including key evidence, interpretation, and any other related partial findings.

- What domain knowledge is essential to facilitate handoff in public avalanche forecasting?
- What knowledge structures can facilitate the capture of knowledge during analysis and how?
- What knowledge structures can facilitate the communication of analysis and how?

8.3. Research Approach

In this study, I employ a combination of participatory design and diary study approaches to investigate how domain knowledge may be structured and implemented in a visual analytics system to better facilitate the capture and communication of analysis in asynchronous collaboration. Participatory design involves actively engaging end-users in the design process, allowing participants to themselves create artefacts. It has often been employed in the context of evaluation in visualization, as it enables the development of visualization tools that closely align with user needs and preferences (T. Isenberg et al., 2013; Lam et al., 2012). Diary studies, on the other hand, involve collecting self-reported data from participants over time, providing valuable insights into their experiences and behaviours (J. Lazar et al., 2017). Often, this involves a log or journal capturing data throughout tasks as well as reflective interviews or questionnaires. Further, the artefacts designed and used within such studies can serve as probes to better understand the potential functionality and design criteria of a tool that does not yet exist. Together, this set of methods is well suited to the progressive and iterative process of exploring design solutions in an ecologically valid manner.

This study is divided into three parts, each conducted during operational forecasting work. Given the exploratory nature of this study and its evolving methodology, the approaches in Parts B and C were shaped by findings in Part A. The initial part, Part A, had two objectives. The first objective of Part A was to identify common categories of information forecasters highlighted in the exercise with the intention of repurposing this structure to more easily capture and organize hand-off materials during analysis. Part C implements this structure in a dedicated knowledge capture tool and investigates how it serves forecasters during their workday. The second objective of Part A was to encourage forecasters to imagine better ways to capture, organize, and communicate hand-off materials. This informed the investigation in Part B, where I examined the operation of one such mechanism in a real-world hand-off scenario.

In all three parts of this study, I employ the diary study method to prompt forecasters to reflect on what is important to capture and communicate while they are working. This approach serves to both inspire design ideation about an as-yet non-existent hand-off support tool and mimic its functionality. This allows forecasters to experience and evaluate the challenges and benefits of using such a tool. This is critical as many potential insights may only be realized through the practice of work, rather than post-hoc reflection. A variety of methods function as “diaries” throughout the three parts of this study. These include think-aloud recordings, written notes, screen-captured images and recordings, sketched diagrams, and questionnaire answers.

8.3.1. Analysis

Research materials such as transcripts, artefacts, and video recordings were reviewed at each stage of research to extract key themes and findings. Given my role in this design process, they are largely a reflection of my own understanding. However, at each stage, I reviewed these themes and findings with other forecasters to check and validate my understanding. As a final step, materials from all three stages were reviewed to synthesize common themes from throughout the study.

8.4. Part A. Identifying Structures & Mechanisms

8.4.1. Objective and Scope

The objectives of this part of the study were to:

1. Identify the types of information captured with the intention of reusing this as a schematization mechanism.
2. Provoke forecasters to reflect on what is important and relevant to share during their workday to help them become aware of their evidential reasoning.
3. Invite forecasters to suggest ways to structure and package materials for collaboration and handoff.

8.4.2. Participants and Procedure

Three forecasters (P2, P3, P11), participated in this stage of the study (Table 9). Forecasters were instructed to record their screens throughout the workday. In addition, they were asked to capture information they deemed important for their own analyses and for hand-off. It was up to their own discretion as to how to do this, though it was recommended that they vocalize their thoughts about such information during the screen recording akin to a think-aloud protocol.

At a later point in time, forecasters were debriefed using video recordings as a reference. Interviews lasted between 30 and 90 minutes. Interviews centered around the kinds of information that forecasters found important and potential design strategies for tools they imagine could be useful to improve hand-off.

Table 9 Participant table for Part A of Study 4

ID	Public Avalanche Forecasting Experience	Background
2	N/A	Marketing, Communications
3	4+	Geological Engineer
11	N/A	Engineering, Natural Science

8.4.3. Data Collection

Data collected include recordings and transcripts from work sessions, debrief interviews, notes the forecasters took and email correspondences between us.

8.4.4. Findings

Common Themes for Structure

Data gathered throughout Part A revealed common categories of information that forecasters deemed important for their analysis and to share with others. P2 spoke over screen-captured video recordings to highlight information they deemed important to carry forward to others on subsequent days. P11 took notes over several days discussing what information is relevant for their specific handover. P3, by contrast, used their workday as inspiration to reflect on and abstract the types of information that would be relevant to carry forward.

Key Evidence

A core topic of information discussed was key or "notable" evidence that influenced forecasters understanding of avalanche conditions and ultimately their assessment decisions. Evidence came in various forms. Often forecasters highlighted individual reports, such as those describing snowpack structure, weather conditions, or observations of avalanches. In Figure 19, P2 captured an image from a report revealing salient visual cues that aid prediction. A "glazed" bed surface and how the crown line of the avalanche "wrapped around" contours of the terrain provided a predictive indication that avalanches could propagate to much larger sizes than those being reported at the time.

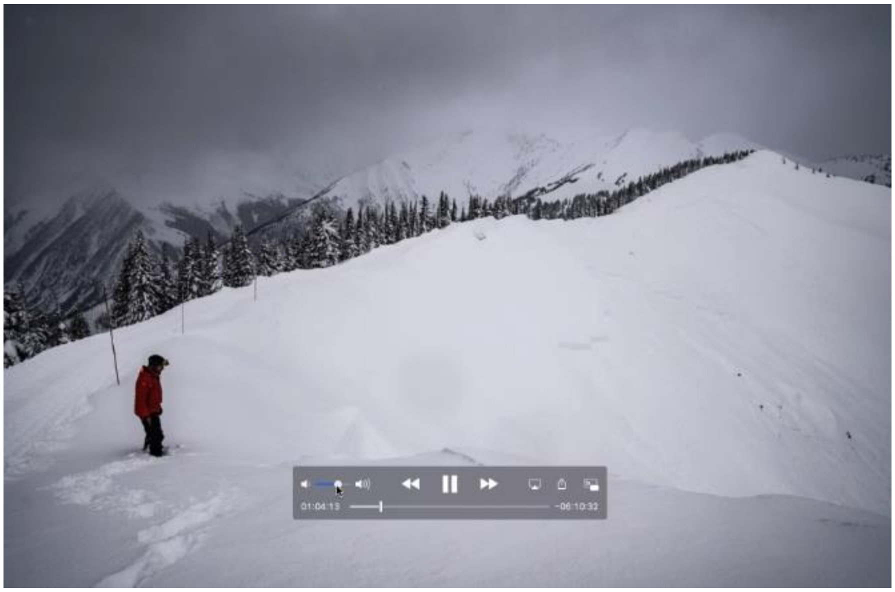


Figure 19 **Evidence captured by P2 that a crust layer may propagate to produce larger avalanches than are being reported currently. P2 noting a "glazed" surface.**

P3 highlighted that understanding which numerical weather prediction models were used in forecasts and how they were interpreted would be a useful form of evidence that is currently rarely discussed.

> I'd say [weather model interpretation] is rarely passed on to the next forecaster right now; that valuable information is usually lost. (P3)

Meanwhile, P11 noted higher-level trends and spatial patterns in data as key pieces of evidence informing their analysis and assessment decisions.

> Tracking past 14 days of avalanche activity (InfoEx and MINs) to determine whether deep persistent problem can be removed or localized (P11)

Hunches

Although they never explicitly labelled these as hunches, forecasters recorded instances where they were developing suspicions backed by little, weak, or difficult-to-articulate evidence. For example, P2 captured an instance where they suspected a snowstorm's trajectory might shift its course to impact their forecast region. This hunch was based on observed trends in the hourly model updates of a numerical weather prediction system (Fig. 20).

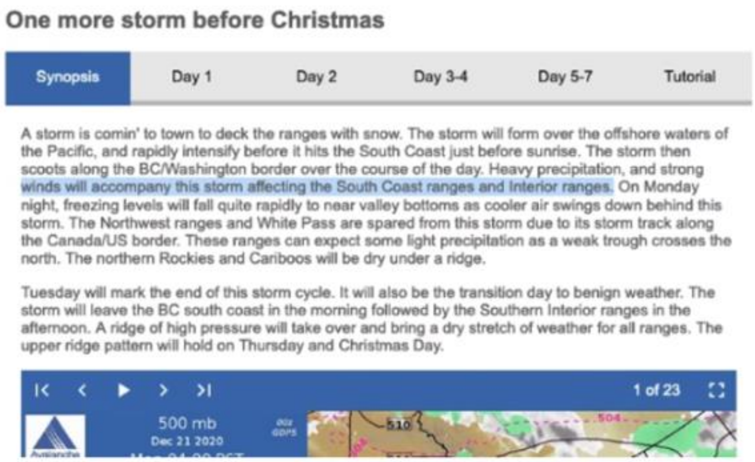


Figure 20 **P2 reasoning through weather model interpretations and developing a hunch based on hourly weather models updates that an incoming weather system is likely to impact their forecast region.**

Meanwhile, P11 picked up on a hunch represented by the prior forecaster's danger rating trend.

> Previous [forecaster's] danger rating trend cued me into their suspicion danger would drop quickly post storm. (P11)

Uncertainty

In our prior research, we observed forecasters often discussing their **uncertainties**. These included gaps in understanding that guide information search or irreducible factors such as weather forecast model uncertainty. We also observed forecasters capturing this information during this study. P11 captured uncertainties multiple times in their notes.

> Main uncertainty: I am assuming there won't be much in terms of new wind slab formation over the weekend (P11)

Decision Rationale

All three forecasters highlighted the importance of providing a rationale for assessment decisions, particularly involving changes to the bulletin that are not obvious or would require additional analysis to understand why an assessment decision might have been made.

> I provide information on why I removed [December weak layer] from the [bulletin] just so that the next person… has some validation as to why it's not there. (P3)

We observed that the rationale forecasters provided for assessment decision elements generally included mentions of key evidence, hunches, uncertainties, and their relationships formulating the explanation.

> Dropped [the danger] rating from MLL to LLL after confirming no avalanche activity after a few days of clear [weather] allowing alpine [observations] (P11)

P3 emphasized the importance of sharing specific facts and evidence to provide enough context for the oncoming forecaster to make sense of the situation and not have to engage in redundant analysis attempting to reconstruct the prior forecaster's assessment decisions.

> If we just say 'I think when slabs are going to continue, and this weak layer is still a problem' without that factual data… that requires the next person to find that factual data to confirm… which is a loss of time and… effort… the previous person would have already gone through those efforts… tracking data is important, and not just the interpretation. (P3)

Persistence of Information

All three forecasters stressed the problem of information persistence. They explained that there is no indication of when documented information such as weak layers or weather summaries have been updated, leaving such information vulnerable to becoming outdated. P11 argued that meta-data such as the last time a piece of information was updated could help coordinate efforts by directing forecasters to focus on determining the relevance of any piece of documented information. Reflecting on his own experience forecasting that day, P2 discussed how explicit directives in hand-off artefacts could direct attention toward work that needs to be done and help coordinate work between forecasters.

> A screenshot of that snow profile, put a red circle around… different weak layers [to indicate] 'maybe you can sort this out' (P2)

Challenges for Schematization

Reflecting on potential mechanisms to schematize knowledge, P11 pointed out that forecasters often invent or use descriptive phrases that evoke understanding based on connotations and shared tacit knowledge. These are important communication vehicles but are not captured in formal standards. For example, P11 defined a “rockies moderate” as a situation where there is a shallow snowpack with many unstable persistent weak layers but few avalanches or “popcorn avalanches” as the last few avalanches that have a latency for being triggered after the primary avalanche cycle, drawing an analogy to the last few kernels of popcorn popping in a bag in the microwave.

P3, on the other hand, pointed out how the use of shorthand phrases can also lead to inefficiencies. They point out that vague words like “recent” to describe events can be very misleading. This word obfuscates the timeline and sequence of events forecasters need to understand current avalanche conditions, particularly when it is carried forward in documentation to future days of work.

Representative Snowprofile for Hand-off

P2 and P3 both suggested using a representative snow profile as a way of organizing hand-off information in a simple and easy-to-understand manner. Such snowprofiles are commonly used in guiding operations to maintain a mental model of snowpack structure and relevant risk considerations. P2 and P3 suggested this could be an effective way of associating interpretation and evidence, and linking it to specific weak layers within the snowpack. In the following section, I provide further background on representative snowprofiles and present findings from an investigation of how such a representative snowprofile may be adapted to support hand-off.

8.5. Part B. Snowprofile Handoff

Representative snowprofiles, common in mountain guiding operations, are used to capture the forecaster’s mental model of snowpack stratigraphy and summarize key aspects of the snowpack to focus on (Canadian Avalanche Association, 1995). These are often recorded in text tables or illustrated in diagrams showing the spatial and physical properties of snow layers relevant to risk assessment (Fig. 21). The y-axis of such diagrams shows the depth of layers which are represented by rectangles. The x-axis represents the hardness of such layers, while the thickness of each rectangle represents the thickness of each layer.

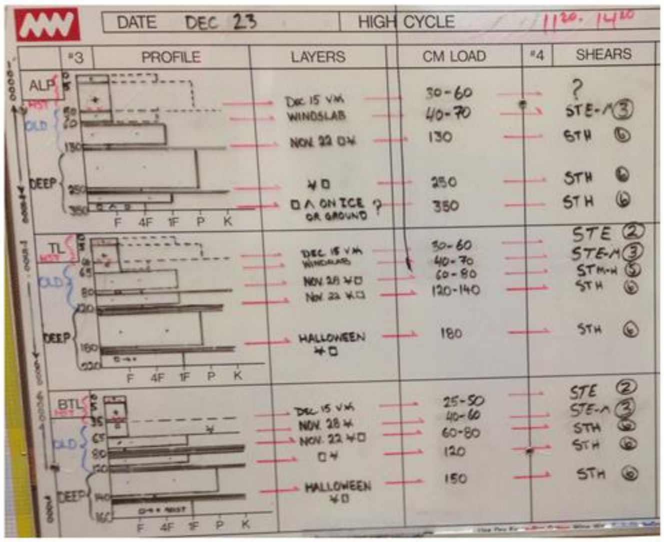


Figure 21 An example representative snowpack summary illustration.
This example was drawn on 23 December 2017 by an avalanche forecaster at Mike Wiegele Helicopter Skiing in Canada (photo: Mike Wiegele Helicopter Skiing)

8.5.1. Objectives and Scope

The purpose of this study was to explore how representative snowprofiles might be repurposed to capture, organize, and communicate hand-off materials in avalanche forecasting. The objectives of this study were to:

1. Observe how a representative snowprofile serves to facilitate hand-off in practice.
2. Invite forecasters to reflect on the utility of this approach.

8.5.2. Participants and Procedure

Two forecasters (P11 & P3) took part in this portion of the study (Table 10). Throughout their workday, one forecaster (P3) created a static digital representative snowprofile diagram, which served as an organizing structure to capture relevant information for hand-off to the oncoming forecaster. It was entirely at their discretion as to how to design this diagram. The receiving forecaster (P11) then used this snowprofile as the starting point for their workday and preceding the review of any other handoff materials that might be available. Both forecasters documented their workday using screen recordings and participated in a 60-minute unstructured debriefing interview.

Table 10 Participant table for Part B of Study 4

ID	Public Avalanche Forecasting Experience	Background
3	4+	Geological Engineer
11	N/A	Engineering, Natural Science

8.5.3. Data Collection

Data collected in this study included recordings and transcripts from the screen recording sessions and debrief interviews. Additionally, the snow profile artifact itself was also considered an essential piece of collected data.

8.5.4. Findings

The authoring forecaster (P3) generated a snowprofile diagram (Fig. 22) annotating layers of the snowpack with notes they deemed relevant to communicate to the oncoming forecaster (P11). This diagram expands on traditional representative snowprofiles by including notes discussing various aspects of P3's reasoning, use of evidence, speculation, and assessment decisions. I note how this visual metaphor effectively captures a variety of information types, however, some **did not fit the structure** and had to be added at the end of the document in the form of freeform text and an image. This underscores how such a structure does not comprehensively account for all relevant information.

> Some freeform text at the bottom that didn't really fit well into the snow profile. And this is pretty common… (P3).

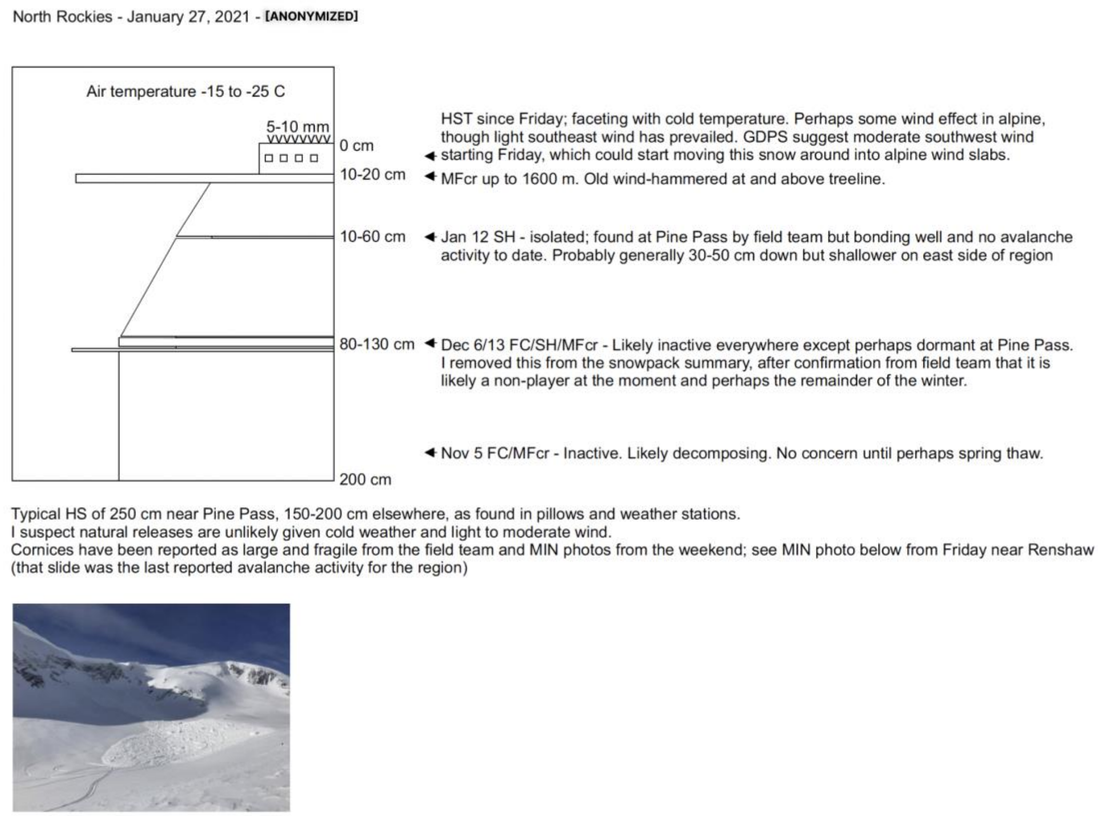


Figure 22 **The snowprofile P3 created and annotated thoughts, impressions, and references to evidence.**

Authoring Forecaster Reflections

While creating the snowprofile, P3 found they could **recognize** important information but could not immediately **articulate** its significance. They found that this approach proved helpful for writing more thorough hand-off notes.

> It took me… a few minutes to realize… sometimes I would say something, and then five minutes later… 'Oh, well, that's why it's important'… And then I would… go back to it… a little bit of a different mindset… But I found it very valuable… usually you only have like five minutes at the very end of your shift to do it… this process really helped me kind of hone-in on what is actually relevant to pass on [and] that'll improve my handover notes (P3)

P3 once again highlighted that **persistence** of information should be a key consideration in designing a production version of such a snowprofile hand-off tool. They discuss how the recency of information can serve as an important indicator for understanding the relative importance of information or whether it requires further investigation. This could serve to help coordinate work between forecasters.

> If you could sort of scroll back to what, what the notes were, like, three or four or five days ago to see what changes were made?... stagnant old data, versus what was actually modified by the most recent forecaster, which is probably the most relevant data. I think that that would be important to highlight. (P3)

P3 also reiterated the importance of providing a **decision rationale** in assessment decisions.

> I provide information on why I removed [December weak layer] from the [bulletin] just so that the next person… have some validation as to why it's not there. (P3)

They stressed how providing evidential details aids the coordination of work between forecasters by eliminating the need to search for such evidence.

> If we just say 'I think when slabs are going to continue, and this weak layer is still a problem' without that factual data… that requires the next person to find that factual data to confirm… which is a loss of time and… effort… the previous person would have already gone through those efforts… (P3)

Receiving Forecaster Reflections

P11, the receiving forecaster, found the visual format of information easy to understand found it improved their ability to **coordinate and plan their workday**. The summary of evidence provided cues where work needed to focus.

> That was the last avalanche from a week ago… it was immediately off my whole workflow… I wasn't gonna spend much time looking at weather stations or filtering through old InfoEx… this kind of shows that work wouldn't be relevant to the current conditions. (P11)

They found it effectively captured a lot of the broader **context** within which to weigh evidence. He explored how this would be very valuable in circumstances of high uncertainty from weather or snowpack structure.

> [It] gave me some context to kind of weigh that evidence against… [when there is] big uncertainty about either the weather forecast, or the snowpack structure itself… that's probably where there'd be more value in notes like this. (P11)

In the recording, it is evident that P11 formed a **hunch** about a weak layer that P3 had added to the snowprofile and decided to focus work on this weak layer.

> Sounds questionable whether or not that will be an avalanche problem. … I kind of want to look around at some of my own sources to see if I can find any more info on that. (P11)

In the later interview, P11 reflected how they directed their attention to the item with the most uncertainty, forming their suspicions using information communicated through the visual format of the snowprofile. They read the depth of the layer in the diagram as a proxy for age and in turn the likely amount of effort that had been spent investigating this layer.

> The main thing that seemed like would be dangerous is if that buried weak layer was more reactive than suspected… that one is more shallowly buried, and there's been less opportunity to observe it or collect info… we've probably been making some assumptions about it [and] haven't really been able to collect a lot of data to validate that. (P11)

Ultimately, they added a new avalanche problem to the bulletin based on this work. However, as this assessment was based on little evidence, they communicated to the oncoming future forecaster that they should try and focus their efforts testing to see if this assessment was valid, fearing that it would be automatically carried forward and become stagnant and erroneous information. They provided a directive to fellow forecasters to coordinate work and ensure that information did not persist longer than needed.

> I tried to push them to remove the problem. If there was any kind of more compelling evidence to do that, where I felt like I made that decision with quite a bit of uncertainty. (P11)

8.6. Part C. Tagging Partial Findings

8.6.1. Objectives and Scope

In the final portion of this study, I repurposed themes derived in Part A (**Key Evidence, Hunch, Uncertainty, Decision Rationale**) as a way of categorizing, or "tagging", information *during work*. Recognizing that it requires additional effort to articulate findings during analysis, this categorization might reduce some of that strain. Moreover, engaging with material through active reading (Brummett, 2018; Mehta et al., 2017; Walny et al., 2018) is thought to improve comprehension and positively benefit work activities. The objectives of this study were to:

1. Investigate how forecasters found such an approach for capturing relevant information during analysis.
2. Investigate whether forecasters found benefit in reflecting on gathered materials.

8.6.2. Participants and Procedure

Table 11 Participant table for Part C of Study 4

ID	Public Avalanche Forecasting Experience	Background
3	4+	Geological Engineer
11	N/A	Engineering, Natural Science
18	N/A	Educator

Three forecasters participated in this stage of the study (Table 11), with two of them utilizing the tool for a single day, while the third utilized it for four consecutive days. Themes identified in Part A were transformed into "tags" – labels or categories for captured information. I did not expect these tags to be comprehensive or to fully represent everything talked about, but simply as a starting point to test how capturing relevant insights during analysis might serve the forecasters in their daily work.

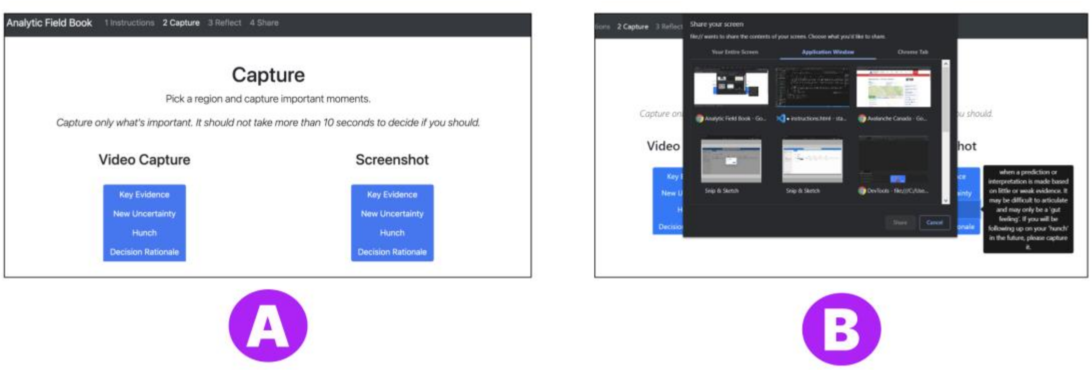


Figure 23 Images of tool used for capturing and "tagging" evidence using screenshots and screen recordings. A) Several buttons allow forecasters to select a tag and whether they would like to capture video or a screenshot. B) This tool uses the screen capture api for modern web browsers.

These tags were incorporated into a browser-based tool allowing forecasters to record their screen and explain their insight or capture a screenshot and write notes in text (Figure 23). Forecasters were instructed to follow a two-step process using this tool. First, forecasters were asked to *capture* their daily work using these tags. Collected media were stored in memory. At the end of the workday, forecasters were then instructed to *reflect* on the information they captured using this tool (Fig. 24). As the focus was primarily on how this information would support forecasters in their daily work, forecasters were not instructed to use these materials directly for hand-off.

Following these work sessions, forecasters were debriefed through email and unrecorded video conference conversations.

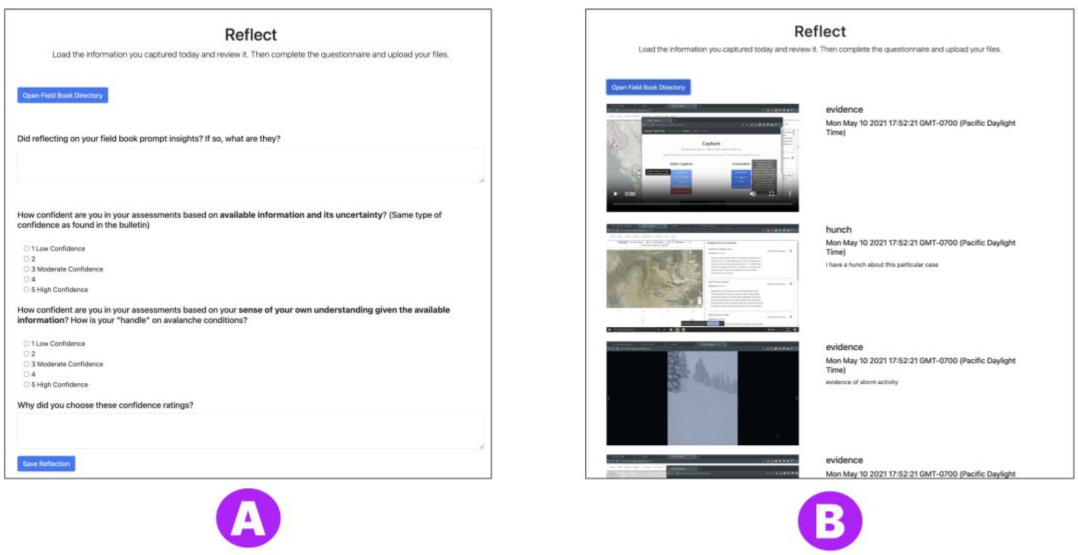


Figure 24 **Images of the interface where forecasters review gathered evidence and their comments and log their reflections. A) A reflection questionnaire. B) An example with captured evidence loaded for reflection.**

8.6.3. Data Collection

Data collected included materials captured using this tool, reflections captured in answers to the questionnaire, and notes taken during debriefing conversations.

8.6.4. Findings

Easy to Recognize, Difficult to Express

Overall, all three forecasters found tagging and capturing relevant information while working relatively easy. However, while forecasters were easily able to *recognize* important information, they sometimes found articulating its significance difficult. P18 found it difficult to categorize the information they were capturing using screen recordings. While they found it easy to recognize which information was important to capture, they found decisions on how to categorize information in the moment more difficult.

> It was all very easy to use. Sometimes I wasn't sure what [tag] to capture it with (P18)

In conversations following the study, P11 and P3 discussed how they found it much easier to identify uncertainties and key evidence than hunches. I speculate that this may be because 'uncertainties' and 'key evidence' are familiar constructs forecasters regularly use in discussions whereas the term and construct of 'hunches' is not. It is a formalism we introduced through our study.

There Isn't Always Something to Capture

P11 reflected how some forecasting days do not involve much change in conditions from prior days and do not yield new insights.

> Today was largely a carry forward day, with most of my analysis simply confirming the previous fxers analysis. Nothing contradictory. Most of my day was spent searching for any counter-evidence, but did not find anything. (P11)

In conversations after the work sessions, they argued that they should be ways to explicitly indicate that no significant changes occurred from prior days and that there are no new insights.

Reflection Aids Work Planning

The process of reflection after a day of gathering notes helped P11 plan their upcoming following workday.

> …what I'm flagging here is things that will prepare me to take a deep dive analysis tomorrow. Given the sparse observations and difficulty knowing the impact of warming as highlighted by these screen captures, I think my forecast tomorrow will be written to communicate uncertainty rather and be a less prescriptive forecast. (P11)

8.7. Discussion

The objective of this study was to explore how domain knowledge may be used to scaffold the capture and organization of hand-off materials during rather than after analysis with the intent of reducing effort and streamlining communication and coordination of work between forecasters. Common themes that shed light on challenges, needs, and opportunities for visual analytics tool designs emerged across all three studies.

First, while forecasters found it **easy to recognize** important information, they found it **difficult to immediately articulate** why it is important. Nevertheless, they found the exercise of reflection beneficial to their work. Second, forecasters repeatedly stressed that any system designed for hand-off will need mechanisms to manage the **persistence** of information in the system. As the state of avalanche conditions is evolving, forecasters must bring knowledge artefacts into a state matching their own mental model formulated through data analysis. Finally, the structure provided by both "tags" in Part C and the spatial layout of the representative snowprofile in Part B, made it easier for forecasters to organize and reflect on work done and in turn **coordinate** work efforts. However, neither of these schematization mechanisms comprehensively captured everything relevant highlighting the limitations of such formalisms in complex and dynamic analysis applications.

Little work in visual analytics has focused on asynchronous sequential handoff of partial findings (Chen et al., 2011; S. Xu et al., 2018; Zhao et al., 2018) and the existing work has focused on exploratory investigative analyses focusing on emergent findings and largely in lab-based settings. By contrast, this application domain involves monitoring dynamically evolving systems and utilizes standardized formalisms for assessing avalanche hazards and communicating them. The judgment of observations along a set of qualitative scales and dimensions is distinct from applications more commonly studied in exploratory visual analysis. The development and application of categories for data are activities common in real-world analytics applications (Bartram et al., 2021). Thus, while this study comprises a small sample of individuals in a working context that is distinct and unique, the issues of how knowledge is constructed and communicated during analysis are not. In the following sections, I draw on key findings from this study as well as existing literature to explore how visual analytics systems can better support the process of knowledge construction and communication in complex and evolving application domains. Specifically, I explore lessons learned and how these inform potential design solutions for how to better capture, organize & manage, and aid the navigation of the meta-data that is the composite and evolving corpus of hand-off materials.

Capture

One of the key lessons I draw for how to best support the capture of information during analysis is to not be overly prescriptive as this can be an impediment to the work at hand. While forecasters found it easy to **recognize** important information, they could not immediately **articulate** why. Forecasters mentioned struggling to decide on how to categorize captured information, particularly when the category was not already in the common parlance of the organization. This suggests that the capture and structuring of information represent distinct thought processes which are often staggered in analysis. Mechanisms for capturing such information should therefore not force categorization at the same time as capture. Simply marking an item as important and bookmarking it to be explained at a later time may be enough and may better serve the flow of analysis.

Organization & Management

Our study shows that existing formalisms and knowledge structures, including those we developed, could not comprehensively account for all relevant information. Forecasters often used freeform text and colloquial phrases to communicate complex situations using shared tacit knowledge and a shared working environment. This reflects the dynamic and evolving nature of shared knowledge constructions. Just as forecasters continuously evolve new ways to express distinct and specific concepts applied to a particular situation, so too should collaborative tools support the natural dynamics of shared knowledge construction. I, therefore, argue for **schematization mechanisms to be flexible and editable.**

Qualitative data analysis tools provide a glimpse of how constructive knowledge processes can be better supported. “Coding” complex information according to common and emergent themes which are grouped, related, and recategorized as need and understanding evolves is a functionality that may prove useful in gathering and schematizing collaborative materials in avalanche forecasting. For instance, new “tag” categories may need to be invented, evidence could be related to multiple constructs, and the relations themselves may additionally have distinct types.

Another issue relates to managing the persistence of captured meta-data. Avalanche forecasting involves monitoring an evolving system and aligning mental models as well as knowledge artefacts with the state of the system. Carrying irrelevant information into the future leads to misunderstanding and misallocation of time.

The context of captured meta-data, such as its timeliness, helps forecasters determine its relevance. In some cases, a visual analytics system could automatically capture contextual information like timestamps. In Part B, P11 used the depth of a buried weak layer as a proxy for its recency and a need for further investigation to determine its relevance. The visual and spatial layout of the snowprofile allowed the forecaster to make this determination. in other instances, forecasters may need to direct their attention more deliberately to questionably relevant information. P11 felt the need to explicitly direct the oncoming future forecaster to evaluate the assessment decisions made based on this same weak layer. Similarly, P2 suggested visually annotating snowprofiles to direct attention toward instabilities requiring more work to coordinate efforts between collaborators.

As the outcome of such investigations involves updates such as the addition or removal of information, these examples point to a need for **passive and active mechanisms for managing the persistence of meta-data.** Passive mechanisms might include visually representing how recently an element has been updated and by whom. This, along with change histories or interaction logs to reconstruct past analysis activities, is a common approach for increasing **awareness** in virtual collaborative environments that could help address issues of persistence management (Drouhard et al., 2017; Marriott et al., 2018; Wu et al., 2013; Zhao et al., 2018). In addition, heatmaps of how frequently certain information was reviewed or worked on could provide a useful cue of where work has been focused and where work remains. Such approaches have been used to facilitate software learning (Matejka et al., 2013)

However, given the heavily context-laden and complex nature of this work, much of what is relevant depends on the forecaster. They determine what information needs further analysis, what may or may not be relevant, and what needs to stay or be removed. Providing a mechanism to "flag" information that needs to be checked and direct attention to may be an effective way for individual forecasters to bookmark items they need to follow-up on or to coordinate work that is in a partial state of completion with other forecasters.

Navigation

Navigating meta-data presents another challenge for visual analytics tools to address. While this study only tested one example of handoff using a snowprofile, our findings do provide an indication of what may be helpful. The representative snowprofile supports various inferences, providing an effective spatial metaphor, and organizing meta-data around the primary focus of forecasters' work: buried weak layer instabilities. As useful as this structure proved to be in our study, I note how most discussion, thoughts, and efforts are not organized around specific named weak layers, but assessment decisions and changes to the bulletin. The rationale for a decision appears to be more common and incorporates other discussed content such as evidence, hunches, uncertainties, or other factors that do not neatly fit the ontology we derived. In this sense, changes to assessments provide a 'common path of travel' or a focal point through which to explore other meta-data. I conjecture that using decision rationales as an organizing structure would therefore be easier to interpret and thus navigate and capture. Moreover, in studies 1 and 3, we observed how forecasters iteratively update the bulletin as they work, bringing it to cohere with their current understanding. In this context, the bulletin itself provides a central organizing structure to scaffold captured meta-data and references to contextual materials. Many domains do not regularly use formalisms like avalanche forecasting. However, when searching for effective domain knowledge constructs explicitly to embed in a visual analytics system to support handoff, focusing on topics that are frequently discussed and centrally related to other topics is best.

In this study, I used screen-captured video and text as proxies for annotations and hypermedia that would be presented in some form of centralized shared view. However, a key challenge to address are mechanisms for navigating between centralized views and other views in which annotated data appears. Annotations are traditionally bound to a representation rather than underlying data, presenting technical challenges in showing annotations in different representational contexts with the same data. This is particularly challenging when using aggregations.

In recent work, researchers explore techniques that bind references to underlying data entities within annotation data structures (Badam et al., 2022). This allows annotations to be displayed in a variety of visual contexts irrespective of data granularity or scale. Meanwhile, the researchers studying collaborative support tools in programming documentation utilize “multi-anchoring”, a method to manually specify what annotations refer to so that it may be displayed in multiple relevant contexts (Horvath et al., 2022). These present promising approaches to better enable navigating between relevant representational contexts, such as the original place in which an annotation was made, and how to ensure that annotations are shown in other contexts where they may be relevant.

This raises questions about what types of additional meta-data need to be captured. For instance, the author, the time of annotation, the originating representational context, the application state, related data entities, or other contexts or objects that the author explicitly deems relevant may need to be captured along with the annotation itself.

This also raises questions about how to best represent annotations in data displays. Given the potential richness of captured meta-data and the volume of annotations, dedicated visual representations and interactions will be needed to avoid visual clutter. A common solution is to use separate coordinated displays such as node-link network knowledge graphs, a concept map, or some other visualization of aggregate annotation structure (Chen & Yang, 2013; Zhao et al., 2018). Badam et al. (2022) use visual overlays to indicate clusters of annotations with shared attributes related to specific data marks. Other subtle approaches such as visual linting (Hopkins et al., 2020), “scented” UI elements (Willett et al., 2007), or other ambient visual markers could provide useful indicators for annotations without cluttering the display or distracting analysis.

8.7.2. Limitations

This formative study provides a preliminary glimpse of how domain knowledge could be used to facilitate handoff in domains involving continuous monitoring and risk assessment. However, this study only included one instance of handoff and gathered a limited set of data exploring meta-data capture. Moreover, no fully functional implementing suggested design features was implemented or tested. While this study has provided rich insights into some of the most central issues arising in this domain, it is limited in scope and not comprehensive. Nevertheless, it provides transferable insights about how visual analytics tools could potentially serve to better support collaborative sensemaking and knowledge construction. Further work is needed to refine and examine the understanding this study suggests.

Chapter 9. Conclusion and Future Work

The key lesson of this dissertation is to **design *for* ambiguity, rather than design it *away*.** Often, what this means is capturing and representing the relevant information as it is: partial and incomplete. This provokes ambiguity and invites the appropriate sensemaking approach. Ambiguity can instill discomfort and in doing so it can motivate action and inquiry. This is a facet of work forecasters are familiar with and expect, which is why they grow suspicious when it is absent. Precision can masquerade as accuracy and when risk is involved, this can have dire consequences. Designing for ambiguity means ensuring that representations support the cognitive work that is needed. Rather than optimizing for precise decoding of summary information, which narrows focus and can obfuscate the *relevant* but difficult-to-anticipate critical details, designing for ambiguity entails showing nuance, complexity, incompleteness, and partiality, inviting the observer to construct their own understanding and supporting them in employing their personal knowledge and sensemaking capabilities.

These understandings challenge normative visualization and interaction design guidelines. Summary overviews or salient perceptual encodings organized into strong visual hierarchies draw attention and increase how accurately information is decoded from a display. However, they do not guarantee that the question or framing presented is the right one to consider. Moreover, relying on interactions to access details presumes that relevant lines of inquiry will be apparent to the viewer at higher levels of abstraction. As we observed, this often is not the case and visualizations can serve to hide what is relevant as much as they can reveal it. In complex analyses, nuance can upend how a problem is conceptualized. Access to raw base data to allow an observer to themselves synthesize and comprehend what is relevant is critical but comprises a set of tasks more akin to *reading* symbols rather than *perceiving patterns* in visual displays of data. This **gisting** is currently not well supported or understood in visualization. It demands further, more rigorous investigation.

9.1. Addressing Research Objectives

In the following section I reflect on how this dissertation has addressed its objective to 1) Characterize the challenges and opportunities for tractable visual analytics solutions in the complex risk-based prediction domain of avalanche forecasting; 2) Design and develop targeted visual analytics solutions; and 3) evaluate these targeted visual analytics solutions by discussing key findings, their significance, and the basis for *transferability*.

9.1.1. Problem Characterization

An understanding of the characteristics that define a problem domain is a prerequisite for evaluating the transferability of findings. The defining characteristics of avalanche forecasting that give rise to ambiguity are that it involves a dynamic complex system, time and resource constraints, incomplete data often gathered using a targeted sampling approach thus requiring deep contextual understanding, and a risk prediction and management context where the cost of "missing" important information can have dire consequences. The limitations of incomplete data and constrained resources produce ambiguity in individual sensemaking. Meanwhile, the interactions between organizations, people – each with different needs, goals, knowledge, constraints, and contexts – and computers, produce ambiguity in collaborative sensemaking.

Study 1 (Chapter 5) surfaced ambiguity as a pervasive aspect of avalanche forecasters' work and contributes an abstraction of how ambiguity arises in analysis which serves to point towards potential visual analytics solutions. I characterize ambiguity as arising from **data, analytic process,** and **collaboration & communication**. These categories are not independent. For example, field reports may be framed as a form of collaborative communication. Nevertheless, these levels point towards the nature and role of ambiguity, the related cognitive work, and associated challenges to suggest visual analytics solutions.

Focusing on ambiguity in sensemaking is significant because it marks a departure from some norms in visual analytics and visualization. Like Andrienko et al. (2018), I treat the goal of visual analysis to be some assessment, prediction, or decision, supported by a process of iterative mental model calibration and development, rather than the generation of *insight*. This broadens the scope of abstraction and our goal as system designers and researchers from the *ends* of a primarily data-driven perspective on analysis (an insight), to a *process* both knowledge-driven and data-driven towards some context-dependant end. Rather than focus on the efficiency and effectiveness of decoding specific information, a ‘microcognitive’ perspective, I use a ‘macrocognitive’ lens to treat and investigate how representations come to aid the cognitive work demanded by a problem domain (Smith et al., 2006). This problem characterization is further refined in subsequent studies.

9.1.2. Designing Visual Analytics for Ambiguity

Study 2 (Chapter 6) explores targeted visual analytics design strategies to support the cognitive work demanded by ambiguity in data.

The key lesson from this participatory design study was that conventional visualization approaches relying on numerical summaries can be problematic because they hide the relevant details necessary for sensemaking. Through iterative refinements, designs departed from conventional guidance in several ways. First, rather than aiming to make decoding information as efficient and automatic as possible, my collaborators and I chose a design strategy that **deliberately uses a visual design that is more difficult to decode as a subversive strategy to encourage scrutiny and more careful consideration of the meaning and implications of data.** In doing so, the necessary difficulties of making sense of this data are preserved rather than designed away. Rather than focusing on a strong visual hierarchy utilizing visual encodings with strong perceptual salience, **we flattened the visual hierarchy** and used multiple visual attributes allowing the observer to exert attentional control and broaden the scope of potential alternative inferences that can be made while preserving a disaggregated view of the data. Finally, we included **visual information with no explicit data mapping to express a visual metaphor**. Our packed circle glyphs used a dynamic force-directed layout that carried no inherent meaning associated with underlying data. Instead, it serves **as a visual metaphor intended to remind viewers about the ambiguous nature of these data**. This contribution adds to the literature employing “desirable difficulties” in visualization (Hullman et al., 2011) and representations of qualitative uncertainties (Boukhelifa et al., 2012).

9.1.3. Evaluating Visual Analytics for Ambiguity from Data

The evaluation study (Study 3, Chapter 7) of AvObs and existing analytics tools provided insight as to how the analytic affordances of AvObs and data tables in InfoEx serve to support or impede sensemaking. When forecasters used existing tools in the InfoEx conditions, they engaged in more ambiguous sensemaking tasks than when using AvObs. Considering the significant effort devoted to accessing report details across both toolsets, along with the fact that tooltip interactions to access these details introduced additional effort and fragmentation to the flow of analysis, the observed differences in ambiguous sensemaking between conditions are likely due to differences in how detailed information is accessed and presented. In contrast to the tooltips in AvObs, the data tables in the existing InfoEx system make details available at a glance. In addition, by showing this information simultaneously, data tables allow forecasters to make use of a distinct analytic affordance: the analytic action of **gisting**. This involves distilling the essence of rich and raw complex information through a mix of automatic and deliberate processes. Gisting is, to the best of my knowledge, not well-studied or discussed in visualization, highlighting a need for further investigation.

Study 3 also contributes a novel approach for evaluating sensemaking in visual analytics. To the best of my knowledge, operationalizing the data-frame theory of sensemaking (Klein et al., 2007) to evaluate a visual analytics system in this way has previously not been done. Whereas sensemaking is conventionally evaluated using insight-based methodologies (North, 2006; Saraiya et al., 2005, 2006) which focus on the volume and type of insights gleaned as well as how visualization features may have supported these, the methodology employed in Study 3 focuses more on specific sensemaking processes and how these may relate to analytic actions. The data-frame theory of sensemaking describes a set of processes involving the manipulation of 'frames', explanations that set expectations for data, while observing and seeking data out in the world.

One limitation of sensemaking evaluation methodologies is the degree to which specific sensemaking processes can be associated with specific visualization features in a cause-and-effect manner. This relationship can only be determined holistically in any ecologically valid research setting. That being said, researchers have been able to trace this relationship in very controlled laboratory conditions, using tasks that do not rely heavily on expert knowledge, using very laborious analysis to gain this high level of resolution (Smuc et al., 2009). The novel sensemaking evaluation used in Study 3 has demonstrated value in yielding useful insights about how visualizations may support cognitive work. As the data-frame theory provides a generalizable model of sensemaking, this evaluation approach may be applied or adapted in other application contexts.

9.1.4. Limitations

As with all ethnographically-inspired formative research, this dissertation has limitations regarding reliability and validity which I would like to acknowledge. First, as it is formative and exploratory, the findings in this research are not conclusive or generalizable. In addition, given the long-term nature of this research, the environment in which the study was carried out and participants were changing. External factors, the technologies developed, as well as my role as a researcher, influenced the organization as it influenced me. Participants who engaged in all or most studies were more subject to being influenced through the research process, while those who engaged in one or a few studies only offered limited design insights as could not develop as rich a shared understanding with me as others.

9.1.5. Summary of Contributions

In addressing the above research objectives, I make the following contributions to the field of visual analytics:

a) a case study of visual analytics applied in a real-world complex risk prediction and management domain;

b) a characterization for how ambiguity arises in the domain of avalanche forecasting to inform the design of visual analytics solutions in risk assessment, prediction, monitoring, and collaboration;

c) a preliminary set of explorations of interactive visualization design strategies and the resulting guidelines for tools to support ambiguity in sensemaking;

d) A qualitative framework and method for evaluating visual analytics tools in complex systems derived from cognitive systems engineering;

e) An evaluation of how different representations can serve to enhance or impede ambiguous sensemaking; and

f) a visual analytics system designed to address ambiguity that has seen field deployment.

9.1.6. Reflection

Drawing on my experience of problem-driven visual analytics research, I offer reflections on this form of research. In particular, I discuss how interactions and understanding developed between myself and the forecasters shaped this research, the practical challenges of conducting work in the field, and thoughts about alternative ways in which this research could have been conducted.

A key aspect of this research was the exchange of knowledge between forecasters and me. At the start of this research, forecasters were primarily familiar with geographic information system (GIS) approaches to visualization which are often distinct from many visualization and interaction paradigms like coordinated and linked multi-view displays. It took some time for forecasters to develop a procedural understanding of how to use interaction as part of a process of inquiry, but they now have come to adopt and expect it in the tools they use. At the same time, this process of adoption also served to reveal new needs and in doing so make explicit how forecasters work. This helped me learn how existing visual analytics approaches do and do not serve forecasters' work and what the nature of their work is. By abstracting tasks and data, we were also translating between two approaches for data analysis and sensemaking. Consequently, there was always a risk of tacit knowledge and procedures being lost in the fray. Having access to both old and new tools available during work seems to have helped the forecasters develop a better understanding of the role each set of tools can play. In doing so, forecasters were able to tell me what was missing and what was not quite right during iterative development. This helped me develop a better understanding of the forecaster's work and shed new light on how existing knowledge in visualization research does or does not accommodate this work.

Coordinating this applied form of research and the associated knowledge exchange was not always easy. Beyond delays in accessing proprietary data, which is common in many visualization application projects, planning research around the forecasters' seasonal and often intensive work was challenging. During the summer forecasters were more available to plan and be involved in design efforts. However, the winter forecasting season proved to be the most valuable time for research and design work. This was when forecasters were most immersed in their work and were able to use prototypes with live data in context to provide feedback that would otherwise have been difficult to anticipate at other times. There were bouts of intensive and closely coordinated rapid iterative design with a handful of forecasters that proved to be very productive because relevant prior design decisions and conversations were easily recalled. When too much time passed in between design iterations, it was more difficult to remind and orient forecasters within intermediary design stages. Finally, given the rapid and multi-modal communication that fit the convenience of the forecasters' work schedules, it was difficult to ensure data was gathered throughout the design process. Much of what was discussed was not recorded and had to be reconstructed through notes, e-mail chains, and design prototypes.

The challenges of balancing research and collaborators' needs is common in applied visualization research. Decisions of how and when to capture research data or when design iterations should stop introducing trade-offs for either party's needs. For instance, in the case of the present research, it may have been interesting to further investigate the concept of "desirable difficulty" by investigating alternative designs that varied the level of effortfulness involved in decoding visualizations. The forecasters may have even welcomed this type of investigation. However, given the practical constraints of such applied research, there are limits to the volume and variety of research inquiries that can be made. This is why my research and design decisions focused on maximizing mutual benefits that avoided narrowing too early on questions of research interest but unknown practical value for forecasters.

In hindsight, there are decisions that I could have made to strengthen this body of research. Instrumenting prototyped tools to log interactions would have provided a highly valuable additional data stream to study how prototypes were used in practice. This would enrich the understanding gleaned from the simulation study (Study 3) and shed light on how tool use might have changed as forecasters adopted it. With hindsight, there are other decisions that could have more productively contributed to specific research questions about aspects of our design approach (e.g., relaxed visual hierarchies, desirable difficulty, visual metaphors, or gisting-support). However, I can only now identify these with any confidence after having conducted this foundational research that inspired these questions in the first place. After all, this is formative and exploratory research intended to raise questions rather than definitively answer them.

Throughout this entire project, I was concerned about balancing my research needs and the forecaster's practical needs. Upon reflection, I realize that any applied visualization research, aiming to support a particular domain group, is certain to produce valuable knowledge contributions to the visualization community. This does not always have to involve novel visualization designs or refinements to existing design guidance. Abstracting and characterizing a problem in an applied setting and mapping this to a set of known visual analytics solutions is a valuable contribution on its own. It reinforces existing understanding in a distinct applied setting that provides evidence setting the basis for transferability.

9.2. Design Implications

I discuss design implications for visual analytics systems aiming to address ambiguity drawing from this dissertation as a cohesive body of work.

1. **Do not unilaterally remove cognitive difficulties in problem solving**. A common theme throughout all the studies conducted in this work is that avalanche forecasting involves much more than what is explicit in data, relying on subtle and difficult to anticipate cues in data that prompt alternative interpretations. This is an inherently difficult task. Conventional visualization design guidelines aim to ease the burden and effortfulness of perceiving patterns and outliers in data using summary overviews and precise perceptual encodings. This capitalizes on the automaticity of low-level perceptual processing but can impart a false sense of precision and, due to the limitations of visual encoding, can occlude details relevant to the task at hand. Designs need take care to balance the advantages of visualization and perceptual processing with the difficulties of a particular problem-solving task and domain. Using perceptually weaker encodings that are more effortful to decode or visual metaphors that remind viewers of the imprecision of data are strategies that can support and evoke the cognitive work necessary to make sense of ambiguous data and situations. In the participatory

design study (Study 2), forecasters found these representational strategies better suited to their needs. When presented with designs that employed conventional design strategies, they felt their sensemaking was impeded and expressed concerns about falsely trusting patterns in data. In addition, the evaluation study and subsequent deployment of AvObs (Study 3) provides further support that the chosen design supports ambiguous sensemaking.

2. **Flatten visual hierarchies.** Whereas salient visual features and strong hierarchies' direct attention to specific aspects of the display, reducing the strength or salience of visual hierarchies allows the viewer themselves the agency to re-target attention to different lines of inquiry as needed. In Study 2, this became apparent from the need to balance disaggregated views of data with broader pattern recognition activities involved in hazard assessment. Glyphs are one well-suited option for this as they can capture numerous data attributes and support a variety of ways for the visual system to traverse visual hierarchies (Borgo et al., 2013). Data tables are another option. As minimalist, precision and salience-emphasizing conventions of digital visualizations are at least partially due to the constraints of early computer graphics, historical printed visualizations and cartography making heavy use of text could serve as inspiration for new designs (Brath, 2020). An important consideration to consider is the scale of data. While the data used by Avalanche Canada is highly complex and rich, it involves relatively lower volumes of data than other domains. This scale certainly affects the interactive representational design strategies that are appropriate. However, it is important to note that in applications involving volumes of data that make comprehensive viewing impractical, data workers still return to sub-samples of large datasets for viewing in a raw form (Bartram et al., 2021).

3. **Reduce interaction overhead and facets.** Interactions naturally introduce more overhead and are generally reserved for information that carries less significance to the primary task at hand. When details are of significance, interactions to access such details can introduce additional overhead. In the evaluation of AvObs (Study 3), forecasters dedicated a larger portion of time to reading tooltips than higher-level patterns. While AvObs provides representational advantages for perceiving patterns, feedback from forecasters and observed usage patterns

suggest that importance of details was undervalued and underemphasized. Further, interaction hierarchies slice the data into layers only retrievable through distinct steps. This can fragment analytic inquiry when the information that needs to be considered simultaneously or compared is only available in successive views through deliberate choice. As there are limits to how much information can be shown simultaneously, the challenge here is to identify which information is relevant to consider so that it is available for viewing. As observed in Study 3, the InfoEx showed the relevant information simultaneously, and thus supported the analytic action of gisting. Moreover, the fact that forecasters still choose to use data tables to diagnose errors in weather station data (Study 2), suggests that gisting may play a role here as well.

4. **Emphasize text reading and gisting.** Throughout all the studies it is clear that details like the identity of a report author, the locations traveled to and manner travelled in, or even how data are expressed offer important cues necessary for forecaster's reasoning processes. While the AvObs prototype has proven useful to the forecasters, having been adopted in operations and demonstrating support for both conventional and ambiguous sensemaking around patterns (Study 3), it occluded views of textual data. Text is of high importance in this domain. Forecasters still use the InfoEx for many core tasks for fear of "missing" critical information. Similarly, in Study 2 we learned that after the WxObs tool was deployed, some forecasters return to raw data tables of weather station data because they have developed ways to scan for and diagnose errors in ways they cannot in aggregate form. Analytics tools addressing such challenges and aiming to provide access to raw base data need to present text in ways that optimally supports sequential reading and rapid scanning of symbols. In addition, care needs to be dedicated to how users navigate between symbolic representations and visualizations. Existing work in the active research area of how text functions within a visualization context may provide further guidance (Brath, 2020). Computational text analysis integrated in such displays could offer additional aid. It could help improve search time when a target is known. However, as it is difficult to anticipate the relevance of any given cue in complex situations, text analysis suffers from the same limitations as aggregate visualizations in trying to reduce or simplify complexity. It therefore should therefore be employed carefully

and not aim to automate away the observer and the natural ambiguity that comes with sensemaking in complex analyses.

9.3. Future Work

This dissertation offers several potential avenues for future work. First, a more in-depth and rigorous study of how AvObs is being used in practice is needed. It is unclear how exactly the tool fits within the broader workflows of hazard assessment, the specific tasks the tool is supporting, and how this relates to ambiguous sensemaking. As the prototype is being used by avalanche professionals in various other organizations, it would be worthwhile to investigate whether or how tool-use varies in these contexts. As this tool in some ways departs from traditional approaches, it may be worthwhile to investigate how this tool has changed work practices. Field observations and situated recall methods, such as those used in Part B of Study 2 may be appropriate here. In addition, evaluations that focus more on specific design features such as the packed circle glyphs and supported visual tasks could help to validate whether they function as intended or in some other way. Focusing attention on these features in a Cued-Recall Debrief interview could provide valuable insights. Alternatively, more lab-based methods with simulated tasks focusing on perceptual tasks could be used. However, findings from such studies would be limited in answering questions about ambiguity, as this deals with more abstract cognitive activities that arise from the complexity of naturalistic settings.

Future research could also more rigorously explore gisting and its role in supporting ambiguous sensemaking. Eye tracking methods (Kurzhals et al., 2014) provide a more direct measure of gaze and attention than mouse cursor movements and could be adapted to the evaluation approach I employed in Study 3. This could offer a higher resolution and more valid perspective of gisting in complex information spaces using symbolic representations (text). Such a study will require careful design considerations. Just as text reading is a learned behaviour, I conjecture gisting in complex information spaces also requires expertise. Selecting an appropriate study population that shares familiarity and expertise using a particular representation will be crucial. In addition, eliciting this behaviour may be challenging as it may depend on the individual, the context, and the available data. As forecasters also rely on meteorological charts and maps that make heavy use of dense symbolic representations, these could offer an alternative to data tables as a context in which to study gisting as an analytic action. Meteorologists, as expert users and authors of such visualizations, would themselves also offer a good target study population.

This also raises questions of what other novel visualization designs that support gisting might be. Inspiration may be drawn from historical or contemporary meteorological visualizations (Houze & Houze, 2019) and historical printed maps, charts and diagrams (Brath, 2020). These may also serve to inform design improvements to the AvObs tool. Based on findings in this research adding text representations, either in the form of data tables or some other idiom, would be of benefit. Of course, such changes would also require further evaluation.

Another potential future line of work addresses the use of domain knowledge as a schematization mechanism to support individual and shared knowledge construction during collaborative analyses. Study 4 yielded only preliminary insights that have yet to be validated. As is customary in visualization design study is, this would require further iterative design, prototyping, and evaluation. Future studies should aim to preserve realism as much as possible given the heavily contextual nature of knowledge work, particularly in as complex a domain as avalanche forecasting.

References

Adams, L. (2005). *A Systems Approach to Human Factors and Expert Decision-Making within Canadian Avalanche Phenomena*. Royal Roads University.

Ahrens, C., & Henson, R. (2018). *Meteorology today: An introduction to weather, climate, and the environment. Cengage Learning*.

Andrienko, N., Lammarsch, T., Andrienko, G., Fuchs, G., Keim, D., Miksch, S., & Rind, A. (2018). Viewing Visual Analytics as Model Building. *Computer Graphics Forum*, *37*(6), 275–299. https://doi.org/10.1111/cgf.13324

Ayyub, B. M. (2001). *Elicitation of expert opinions for uncertainty and risks*. CRC Press.

Badam, S. K., Chandrasegaran, S., & Elmqvist, N. (2022). Integrating annotations into multidimensional visual dashboards. *Information Visualization*, *21*(3), 270–284. https://doi.org/10.1177/14738716221079591

Bartram, L., Correll, M., & Tory, M. (2021). Untidy Data: The Unreasonable Effectiveness of Tables. *arXiv:2106.15005 [Cs]*. http://arxiv.org/abs/2106.15005

Beach, L. R. (2009). Narrative thinking and decision making: How the stories we tell ourselves shape our decisions, and vice versa. *Online Publication: Www. LeeRoyBeach. Com*.

Bentley, T., Johnston, L., & von Baggo, K. (2005). Evaluation Using Cued-recall Debrief to Elicit Information About a User's Affective Experiences. *Proceedings of the 17th Australia Conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*, 1–10. http://dl.acm.org/citation.cfm?id=1108368.1108403

Bertin, J. (1983). Semiology of graphics (WJ Berg, Trans.). *Madison, WI: The University of Wisconsin Press.(Original Work Published 1967)*.

Bertini, E., Correll, M., & Franconeri, S. (2020). Why Shouldn't All Charts Be Scatter Plots? Beyond Precision-Driven Visualizations. *2020 IEEE Visualization Conference (VIS)*, 206–210. https://doi.org/10.1109/VIS47514.2020.00048

Beven, K. J., Almeida, S., Aspinall, W. P., Bates, P. D., Blazkova, S., Borgomeo, E., Goda, K., Hall, J. W., Phillips, J. C., Simpson, M., Smith, P. J., Stephenson, D. B., Wagener, T., Watson, M., & Wilkins, K. L. (2018). Epistemic uncertainties and natural hazard risk assessment. 1. A review of different natural hazard areas. *Natural Hazards and Earth System Sciences*. https://doi.org/10.5194/nhess-2017-250

Borgo, R., Kehrer, J., Chung, D. H., Maguire, E., Laramee, R. S., Hauser, H., Ward, M., & Chen, M. (2013). Glyph-based Visualization: Foundations, Design Guidelines, Techniques and Applications. *Eurographics (State of the Art Reports)*, 39–63.

Boukhelifa, N., Bezerianos, A., Isenberg, T., & Fekete, J. (2012). Evaluating Sketchiness as a Visual Variable for the Depiction of Qualitative Uncertainty. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2769–2778. https://doi.org/10.1109/TVCG.2012.220

Boukhelifa, N., Perrin, M.-E., Huron, S., & Eagan, J. (2017). How Data Workers Cope with Uncertainty: A Task Characterisation Study. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3645–3656. https://doi.org/10.1145/3025453.3025738

Brath, R. (2020). *Visualizing with Text*. CRC Press.

Braun, V., & Clarke, V. (2012). *Thematic analysis.* https://doi.org/10.1037/13620-004

Brehmer, M., & Munzner, T. (2013). A Multi-Level Typology of Abstract Visualization Tasks. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2376–2385. https://doi.org/10.1109/TVCG.2013.124

Brummett, B. (2018). *Techniques of close reading*. Sage Publications.

Bruner, J. S. (2009). *Actual Minds, Possible Worlds*. Harvard University Press.

Campbell, Gould, Haegeli, & Conger. (2016). *Technical aspects of snow avalanche risk management: Resources and guidelines for avalanche practitioners in Canada.*

Canadian Avalanche Association. (1995). *Observation guidelines and recording standards for weather, snowpack and avalanches*. Canadian Avalanche Assoc.

Canadian Avalanche Association. (1995). *Observation guidelines and recording standards for weather, snowpack and avalanches*. [Revelstoke, BC]: Canadian Avalanche Assoc.

Card, S., Mackinlay, J., & Shneiderman, B. (1999). Readings in Information Visualization: Using Vision To Think. In *Information Visualization—IVS*.

Carpendale, S. (2008). Evaluating Information Visualizations. In A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North (Eds.), *Information Visualization* (Vol. 4950, pp. 19–45). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-70956-5_2

Chen, Y., Alsakran, J., Barlowe, S., Yang, J., & Zhao, Y. (2011). Supporting effective common ground construction in Asynchronous Collaborative Visual Analytics. *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 101–110. https://doi.org/10.1109/VAST.2011.6102447

Chen, Y., & Yang, J. (2013). ManyInsights: A visual analytics approach to supporting effective insight management. *Tsinghua Science and Technology*, *18*(2), 171–181. https://doi.org/10.1109/TST.2013.6509100

Cheong, L., Kinkeldey, C., Burfurd, I., Bleisch, S., & Duckham, M. (2020). Evaluating the impact of visualization of risk upon emergency route-planning. *International Journal of Geographical Information Science*, *34*(5), 1022–1050. https://doi.org/10.1080/13658816.2019.1701677

Chi, M. T. (2013). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In *Advances in instructional psychology* (pp. 161–238). Routledge.

Chung, H., Yang, S., Massjouni, N., Andrews, C., Kanna, R., & North, C. (2010). VizCept: Supporting synchronous collaboration for constructing visualizations in intelligence analysis. *2010 IEEE Symposium on Visual Analytics Science and Technology*, 107–114. https://doi.org/10.1109/VAST.2010.5652932

Clark, T. (2019). *Exploring the link between the conceptual model of avalanche hazard and the North American public avalanche danger scale*.

Cleveland, W. S., & McGill, R. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, *79*(387), 531–554. https://doi.org/10.1080/01621459.1984.10478080

Cook, K. A., & Thomas, J. J. (2005). *Illuminating the path: The research and development agenda for visual analytics*. Pacific Northwest National Lab.(PNNL), Richland, WA (United States).

Cooke, L. (2010). Assessing Concurrent Think-Aloud Protocol as a Usability Test Method: A Technical Communication Approach. *IEEE Transactions on Professional Communication*, *53*(3), 202–215. https://doi.org/10.1109/TPC.2010.2052859

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. MIT Press.

Diaconis, P. (2006). Theories of data analysis: From magical thinking through classical statistics. *Exploring Data Tables, Trends, and Shapes*, 1–36.

Douven, I. (2017). Abduction. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2017/entries/abduction/

Dow, S., MacIntyre, B., Lee, J., Oezbek, C., Bolter, J. D., & Gandy, M. (2005). Wizard of Oz support throughout an iterative design process. *IEEE Pervasive Computing*, *4*(4), 18–26. https://doi.org/10.1109/MPRV.2005.93

Drouhard, M., Chen, N.-C., Suh, J., Kocielnik, R., Peña-Araya, V., Cen, K., Xiangyi Zheng, & Aragon, C. R. (2017). Aeonium: Visual analytics to support collaborative qualitative coding. *2017 IEEE Pacific Visualization Symposium (PacificVis)*, 220–229. https://doi.org/10.1109/PACIFICVIS.2017.8031598

Elfenbein, A. (2018). *The Gist of Reading*. Stanford University Press. http://ebookcentral.proquest.com/lib/sfu-ebooks/detail.action?docID=5178043

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. *The Quarterly Journal of Economics*, *75*(4), 643–669. JSTOR. https://doi.org/10.2307/1884324

Federico, P., Wagner, M., Rind, A., Amor-Amorós, A., Miksch, S., & Aigner, W. (2017). *The Role of Explicit Knowledge: A Conceptual Model of Knowledge-Assisted Visual Analytics*. 12. https://doi.org/10.1109/VAST.2017.8585498

Fischer, A. P. (2011). The Measurement Factors in Estimating Snowfall Derived from Snow Cover Surfaces Using Acoustic Snow Depth Sensors. *Journal of Applied Meteorology and Climatology*, *50*(3), 681–699. https://doi.org/10.1175/2010JAMC2408.1

Fischer, F., Bruhn, J., Gräsel, C., & Mandl, H. (2002). Fostering collaborative knowledge construction with visualization tools. *Learning and Instruction*, *12*(2), 213–232. https://doi.org/10.1016/S0959-4752(01)00005-6

Gamino, J., Chapman, S., Hull, E., & Lyon, R. (2010). Effects of Higher-Order Cognitive Strategy Training on Gist-Reasoning and Fact-Learning in Adolescents. *Frontiers in Psychology*, *1*. https://doi.org/10.3389/fpsyg.2010.00188

Gao, T., Dontcheva, M., Adar, E., Liu, Z., & Karahalios, K. G. (2015). DataTone: Managing Ambiguity in Natural Language Interfaces for Data Visualization. *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, 489–500. https://doi.org/10.1145/2807442.2807478

Gaver, W. (2012). What should we expect from research through design? *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 937–946. https://doi.org/10.1145/2207676.2208538

Geisler, C., & Swarts, J. (2019). *Coding streams of language: Techniques for the systematic coding of text, talk, and other verbal data*. WAC Clearinghouse Ft. Collins, CO.

Goyal, N., Leshed, G., Cosley, D., & Fussell, S. R. (2014). Effects of Implicit Sharing in Collaborative Analysis. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 129–138. https://doi.org/10.1145/2556288.2557229

Grabowski, F., & Strzalka, D. (2008). Simple, complicated and complex systems—The brief introduction. *2008 Conference on Human System Interactions*, 570–573. https://doi.org/10.1109/HSI.2008.4581503

Green, T. M., & Ribarsky, W. (2008). *Using a human cognition model in the creation of collaborative knowledge visualizations* (M. S. Kim, K. Chao, W. J. Tolone, W. Ribarsky, S. I. Balandin, B. Javidi, & S.-I. Tu, Eds.; p. 69830C). https://doi.org/10.1117/12.777189

Haegeli, P., Obad, J., Harrison, B., Murray, B., Engblom, J., & Neufeld, J. (2014). InfoEx™ 3.0—Advancing The Data Analysis Capabilities Of Canada's Diverse

Avalanche Community. *Proceedings of the International Snow Science Workshop*, 8.

Hartson, R., & Pyla, P. (2019). Chapter 19 - Background: Design. In R. Hartson & P. Pyla (Eds.), *The UX Book (Second Edition)* (pp. 397–401). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-805342-3.00019-9

Healey, C. G., & Enns, J. T. (2012). Attention and Visual Memory in Visualization and Computer Graphics. *IEEE Transactions on Visualization and Computer Graphics*, *18*(7), 1170–1188. https://doi.org/10.1109/TVCG.2011.127

Heer, J., & Agrawala, M. (2008). Design Considerations for Collaborative Visual Analytics. *Information Visualization*, *7*(1), 49–62. https://doi.org/10.1057/palgrave.ivs.9500167

Hilligoss, B., & Moffatt-Bruce, S. D. (2014). The limits of checklists: Handoff and narrative thinking. *BMJ Quality & Safety*, *23*(7), 528–533. https://doi.org/10.1136/bmjqs-2013-002705

Hoffman, R. R., Coffey, J. W., Ford, K. M., & Novak, J. D. (2006). A Method for Eliciting, Preserving, and Sharing the Knowledge of Forecasters. *Weather and Forecasting*, *21*(3), 416–428. https://doi.org/10.1175/WAF927.1

Hoffman, R. R., LaDue, D. S., Mogil, H. M., Roebber, P. J., & Trafton, J. G. (2017). *Minding the Weather: How Expert Forecasters Think*. The MIT Press. http://muse.jhu.edu/book/53003

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC Press.

Hopkins, A. K., Correll, M., & Satyanarayan, A. (2020). VisuaLint: Sketchy In Situ Annotations of Chart Construction Errors. *Computer Graphics Forum*, *39*(3), 219–228. https://doi.org/10.1111/cgf.13975

Hoque, E., Setlur, V., Tory, M., & Dykeman, I. (2018). Applying Pragmatics Principles for Interaction with Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 309–318. https://doi.org/10.1109/TVCG.2017.2744684

Hordowick, H. (2022). *Understanding avalanche problem assessments: A concept mapping study with public avalanche forecasters* [MRM, School for Resource and Environmental Management]. Simon Fraser University.

Horton, S., Nowak, S., & Haegeli, P. (2018). Exploring regional snowpack patterns with gridded models. *Proceedings of International Snow Science Workshop, Innsbruck, Austria*, 1136–1140.

Horton, S., Nowak, S., & Haegeli, P. (2020). Enhancing the operational value of snowpack models with visualization design principles. *Natural Hazards and Earth System Sciences*, *20*(6), 1557–1572. https://doi.org/10.5194/nhess-20-1557-2020

Horvath, A., Liu, M. X., Hendriksen, R., Shannon, C., Paterson, E., Jawad, K., Macvean, A., & Myers, B. A. (2022). Understanding How Programmers Can Use Annotations on Documentation. *CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3491102.3502095

Houze, R. A., & Houze, R. (2019). Cloud and weather symbols in the historic language of weather map plotters. *Bulletin of the American Meteorological Society*, *100*(12), ES423–ES443. https://doi.org/10.1175/BAMS-D-19-0071.1

Hullman, J., Adar, E., & Shah, P. (2011). Benefitting InfoVis with Visual Difficulties. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2213–2222. https://doi.org/10.1109/TVCG.2011.175

International Organization for Standardization. (2009). *ISO 31000: Risk Management: Principles and Guidelines*. ISO.

Isenberg, P., Elmqvist, N., Scholtz, J., Cernea, D., Ma, K.-L., & Hagen, H. (2011). Collaborative visualization: Definition, challenges, and research agenda. *Information Visualization*, *10*(4), 310–326. https://doi.org/10.1177/1473871611412817

Isenberg, T., Isenberg, P., Jian Chen, Sedlmair, M., & Moller, T. (2013). A Systematic Review on the Practice of Evaluating Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2818–2827. https://doi.org/10.1109/TVCG.2013.126

Jackson, H. J. (2001). *Marginalia: Readers writing in books*. Yale University Press.

Jamieson, B., Haegeli, P., & Statham, G. (2015). *Uncertainty in Snow Avalanche Risk Assessments*. *35*, 701–712. https://doi.org/10.1111/risa.12298

Jänicke, S., Kaur, P., Kuzmicki, P., & Schmidt, J. (2020). Participatory Visualization Design as an Approach to Minimize the Gap between Research and Application. *VisGap@ Eurographics/EuroVis*, 35–42.

Johansen, I. L., & Rausand, M. (2015). Ambiguity in risk assessment. *Safety Science*, *80*, 243–251. https://doi.org/10.1016/j.ssci.2015.07.028

Ju, Y., & You, M. (2018). Developing a gist-extraction typology based on journalistic lead writing: A case of food risk news. *Heliyon*, *4*(8), e00738. https://doi.org/10.1016/j.heliyon.2018.e00738

Kamal, A., Dhakal, P., Javaid, A. Y., Devabhaktuni, V. K., Kaur, D., Zaientz, J., & Marinier, R. (2021). Recent advances and challenges in uncertainty visualization: A survey. *Journal of Visualization*, *24*(5), 861–890. https://doi.org/10.1007/s12650-021-00755-1

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. In S. J. Simoff, M. H. Böhlen, & A. Mazeika (Eds.), *Visual Data Mining: Theory, Techniques and Tools for Visual Analytics* (pp. 76–90). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-71080-6_6

Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., & Melançon, G. (2008). Visual Analytics: Definition, Process, and Challenges. In A. Kerren, J. T. Stasko, J.-D. Fekete, & C. North (Eds.), *Information Visualization: Human-Centered Issues and Perspectives* (pp. 154–175). Springer. https://doi.org/10.1007/978-3-540-70956-5_7

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

Kirsh, D. (2010). Thinking with external representations. *AI & SOCIETY*, *25*(4), 441–454. https://doi.org/10.1007/s00146-010-0272-8

Klassen, K. (2013). What's the problem? A primer on defining avalanche character. In *The Avalanche Journal* (Vol. 105, pp. 10–12).

Klassen, K., Haegeli, P., Consulting, A., Statham, G., & Canada, P. (2013). The Role of Avalanche Character in Public Avalanche Safety Products. *International Snow Science Workshop*, 7.

Klein, G., & Crandall, B. W. (2018). The Role of Mental Simulation in Problem Solving and Decision Making. In *Local Applications of the Ecological Approach to Human-Machine Systems* (Vol. 2, pp. 324–358). CRC Press.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making Sense of Sensemaking 1: Alternative Perspectives. *IEEE Intelligent Systems*, *21*(4), 70–73. https://doi.org/10.1109/MIS.2006.75

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. *Expertise out of Context: Proceedings of the Sixth International Conference on Naturalistic Decision Making*, 113–155.

Klein, G., Pliske, R., Crandall, B., & Woods, D. (2005). Problem detection. *Cognition, Technology & Work*, *7*, 14–28. https://doi.org/10.1007/s10111-004-0166-y

Klein, G., Pliske, R. M., Crandall, B., & Woods, D. (1999). Features of Problem Detection. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *43*(3), 133–137. https://doi.org/10.1177/154193129904300301

Klein, G., & Snowden, D. (2011). Anticipatory Thinking. *Informed by Knowledge Expert Performance in Complex Situations*. https://doi.org/10.4324/9780203847985

Koffka, K. (1935). *Principles of Gestalt psychology* (Vol. 44). Routledge.

Kurzhals, K., Fisher, B., Burch, M., & Weiskopf, D. (2014). Evaluating visual analytics with eye tracking. *Proceedings of the Fifth Workshop on Beyond Time and Errors: Novel Evaluation Methods for Visualization*, 61–69.

LaChapelle, E. R. (1980). The Fundamental Processes in Conventional Avalanche Forecasting. *Journal of Glaciology*, *26*(94), 75–84.

Lam, H., Bertini, E., Isenberg, P., Plaisant, C., & Carpendale, S. (2012). Empirical Studies in Information Visualization: Seven Scenarios. *IEEE Transactions on Visualization and Computer Graphics*, *18*(9), 1520–1536. https://doi.org/10.1109/TVCG.2011.279

Lazar, B., Trautman, S., Cooperstein, M., Greene, E., & Birkeland, K. (2016). North American Avalanche Danger Scale: Do Backcountry Forecasters Apply It Consistently? *International Snow Science Workshop*, 9.

Lazar, J., Feng, J. H., & Hochheiser, H. (2017). *Research methods in human-computer interaction*. Morgan Kaufmann.

Lin, H., Akbaba, D., Meyer, M., & Lex, A. (2021). Data Hunches: Incorporating Personal Knowledge into Visualizations. *arXiv:2109.07035 [Cs]*. http://arxiv.org/abs/2109.07035

Liu, J., Boukhelifa, N., & Eagan, J. R. (2020). Understanding the Role of Alternatives in Data Analysis Practices. *IEEE Transactions on Visualization and Computer Graphics*, *26*(1), 66–76. https://doi.org/10.1109/TVCG.2019.2934593

Lundquist, J., Hughes, M., Gutmann, E., & Kapnick, S. (2019). Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks. *Bulletin of the American Meteorological Society*, *100*(12), 2473–2490. https://doi.org/10.1175/BAMS-D-19-0001.1

MacEachren, A. M. (2015). Visual Analytics and Uncertainty: Its Not About the Data. *EuroVA@EuroVis*. https://doi.org/10.2312/eurova.20151104

MacEachren, A. M., Robinson, A., Hopper, S., Gardner, S., Murray, R., Gahegan, M., & Hetzler, E. (2005). Visualizing geospatial information uncertainty: What we know and what we need to know. *Cartography and Geographic Information Science*, *32*(3), 139–160. https://doi.org/10.1559/1523040054738936

Mahyar, N., & Tory, M. (2014). Supporting Communication and Coordination in Collaborative Sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, *20*(12), 1633–1642. https://doi.org/10.1109/TVCG.2014.2346573

Maitlis, S., & Christianson, M. (2014). Sensemaking in Organizations: Taking Stock and Moving Forward. *The Academy of Management Annals*, *8*. https://doi.org/10.1080/19416520.2014.873177

Malik, A., Maciejewski, R., Maule, B., & Ebert, D. S. (2011). A visual analytics process for maritime resource allocation and risk assessment. *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 221–230. https://doi.org/10.1109/VAST.2011.6102460

Marriott, K., Schreiber, F., Dwyer, T., Klein, K., Riche, N. H., Itoh, T., Stuerzlinger, W., & Thomas, B. H. (Eds.). (2018). *Immersive Analytics* (Vol. 11190). Springer International Publishing. https://doi.org/10.1007/978-3-030-01388-2

Matejka, J., Grossman, T., & Fitzmaurice, G. (2013). Patina: Dynamic heatmaps for visualizing application usage. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3227–3236. https://doi.org/10.1145/2470654.2466442

Mathisen, A., Horak, T., Klokmose, C. N., Grønbæk, K., & Elmqvist, N. (2019). InsideInsights: Integrating Data-Driven Reporting in Collaborative Visual Analytics. *Computer Graphics Forum*, *38*(3), 649–661. https://doi.org/10.1111/cgf.13717

McClung, D. M. (2002a). The Elements of Applied Avalanche Forecasting, Part I: The Human Issues. *Natural Hazards*, *26*(2), 111–129.

McClung, D. M. (2002b). The Elements of Applied Avalanche Forecasting, Part II: The Physical Issues and the Rules of Applied Avalanche Forecasting. *Natural Hazards*, *26*(2), 131–146.

McClung, D. M., & Schaerer, P. (2006). *The Avalanche Handbook*. The Mountaineers Books.

Mccurdy, N., Gerdes, J., & Meyer, M. (2019). A Framework for Externalizing Implicit Error Using Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 925–935. https://doi.org/10.1109/TVCG.2018.2864913

McNutt, A., Kindlmann, G., & Correll, M. (2020). Surfacing Visualization Mirages. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. https://doi.org/10.1145/3313831.3376420

Mehta, H., Bradley, A., Hancock, M., & Collins, C. (2017). Metatation: Annotation As Implicit Interaction to Bridge Close and Distant Reading. *ACM Trans. Comput.-Hum. Interact.*, *24*(5), 35:1-35:41. https://doi.org/10.1145/3131609

Mekis, E., Donaldson, N., Reid, J., Zucconi, A., Hoover, J., Li, Q., Nitu, R., & Melo, S. (2018). An Overview of Surface-Based Precipitation Observations at

Environment and Climate Change Canada. *Atmosphere-Ocean*, *56*(2), 71–95. https://doi.org/10.1080/07055900.2018.1433627

Morin, S., Horton, S., Techel, F., Bavay, M., Coléou, C., Fierz, C., Gobiet, A., Hagenmuller, P., Lafaysse, M., Ližar, M., Mitterer, C., Monti, F., Müller, K., Olefs, M., Snook, J. S., van Herwijnen, A., & Vionnet, V. (2020). Application of physical snowpack models in support of operational avalanche hazard forecasting: A status report on current implementations and prospects for the future. *Cold Regions Science and Technology*, *170*, 102910. https://doi.org/10.1016/j.coldregions.2019.102910

Mueller, F. “Floyd,” Kethers, S., Alem, L., & Wilkinson, R. (2006). From the Certainty of Information Transfer to the Ambiguity of Intuition. *Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, 63–70. https://doi.org/10.1145/1228175.1228189

Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6). https://doi.org/10.1109/TVCG.2009.111

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*(3), 231–259. https://doi.org/10.1037/0033-295X.84.3.231

North, C. (2006). Toward measuring visualization insight. *IEEE Computer Graphics and Applications*, *26*(3), 6–9. https://doi.org/10.1109/MCG.2006.70

Nowak, S. (2019a). *Designing Visualization Tools for Avalanche Forecasters*.

Nowak, S. (2019b). What Can Visual Analytics Do for Avalanche Forecasting? *The Avalanche Journal*, *120*.

Nowak, S. (2021). *Interactive Visualization for Avalanche Hazard Assessment*.

Nowak, S., Bartram, L., & Schiphorst, T. (2018). A Micro-Phenomenological Lens for Evaluating Narrative Visualization. *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)*, 11–18. https://doi.org/10.1109/BELIV.2018.8634072

Omodei, M. M., & McLennan, J. (1994). Studying Complex Decision Making in Natural Settings: Using a Head-Mounted Video Camera to Study Competitive Orienteering. *Perceptual and Motor Skills*, *79*(3_suppl), 1411–1425. https://doi.org/10.2466/pms.1994.79.3f.1411

Padilla, L. M. (2018). A Case for Cognitive Models in Visualization Research: Position paper. *2018 IEEE Evaluation and Beyond - Methodological Approaches for Visualization (BELIV)*, 69–77. https://doi.org/10.1109/BELIV.2018.8634267

Panagiotidou, G., Vandam, R., Poblome, J., & Vande Moere, A. (2021). Implicit Error, Uncertainty and Confidence in Visualization: An Archaeological Case Study. *IEEE Transactions on Visualization and Computer Graphics*, 1–1. https://doi.org/10.1109/TVCG.2021.3088339

Patterson, E. S. (2008). Structuring flexibility: The potential good, bad and ugly in standardisation of handovers. *Quality and Safety in Health Care*, *17*(1), 4–5. https://doi.org/10.1136/qshc.2007.022772

Patterson, E. S., Roth, E. M., Woods, D. D., Chow, R., & Gomes, J. O. (2004). Handoff strategies in settings with high consequences for failure: Lessons for health care operations. *International Journal for Quality in Health Care*, *16*(2), 125–132. https://doi.org/10.1093/intqhc/mzh026

Patterson, E. S., Watts-Perotti, J., & Woods, D. D. (2017). Voice Loops: Engineering Overhearing to Aid Coordination. In C. P. Nemeth (Ed.), *Improving Healthcare Team Communication* (1st ed., pp. 79–95). CRC Press. https://doi.org/10.1201/9781315588056-6

Patterson, E. S., & Woods, D. D. (2001). Shift changes, updates, and the on-call architecture in space shuttle mission control. *Computer Supported Cooperative Work (CSCW)*, *10*(3–4), 317–346. https://doi.org/10.1023/A:1012705926828

Patterson, E. S., Woods, D. D., Cook, R. I., & Render, M. L. (2007). Collaborative cross-checking to enhance resilience. *Cognition, Technology & Work*, *9*(3), 155–162. https://doi.org/10.1007/s10111-006-0054-8

Pirolli, P., & Card, S. (2005). *The Sensemaking Process and Leverage Points for Analyst Technology as Identified Through Cognitive Task Analysis*. Proceedings of International Conference on Intelligence Analysis.

Pliske, R., Crandall, B., & Klein, G. (2004). Competence in weather forecasting. *Psychological Investigations of Competence in Decision Making*, 40–70.

Prue, B., Jenkins, M., Stern, L. D., & Pfautz, J. (2014). Overcoming Limited Collaboration Channels in Distributed Intelligence Analysis: Visualization Tools and Design Seeds. *2014 IEEE Joint Intelligence and Security Informatics Conference*, 17–24. https://doi.org/10.1109/JISIC.2014.13

Rädle, R., Nouwens, M., Antonsen, K., Eagan, J. R., & Klokmose, C. N. (2017). Codestrates: Literate Computing with Webstrates. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, 715–725. https://doi.org/10.1145/3126594.3126642

Rensink, R. A. (2000). The Dynamic Representation of Scenes. *Visual Cognition*, *7*(1–3), 17–42. https://doi.org/10.1080/135062800394667

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision Making*, *7*(3), 332–359. https://doi.org/10.1017/S1930297500002291

Riche, N. H. (Ed.). (2018). *Data-driven storytelling*. Taylor & Francis/CRC Press.

Rind, A., Aigner, W., Wagner, M., Miksch, S., & Lammarsch, T. (2016). Task Cube: A three-dimensional conceptual space of user tasks in visualization design and evaluation. *Information Visualization*, *15*(4), 288–300. https://doi.org/10.1177/1473871615621602

Rind, A., Wagner, M., & Aigner, W. (2019). Towards a Structural Framework for Explicit Domain Knowledge in Visual Analytics. *arXiv:1908.07752 [Cs]*. http://arxiv.org/abs/1908.07752

Rod, J. K. (2001). How the monosemic graphics go polysemic. *Cartographic Perspectives*, *38*, 7–18. https://doi.org/10.14714/CP38.792

Saraiya, P., North, C., & Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *11*(4), 443–456. https://doi.org/10.1109/TVCG.2005.53

Saraiya, P., North, C., Vy Lam, & Duca, K. A. (2006). An Insight-Based Longitudinal Study of Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics*, *12*(6), 1511–1522. https://doi.org/10.1109/TVCG.2006.85

Schuler, D., & Namioka, A. (1993). *Participatory Design: Principles and Practices*. L. Erlbaum Associates Inc.

Sedlmair, M., Meyer, M., & Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics*, *18*(12), 2431–2440. https://doi.org/10.1109/TVCG.2012.213

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

Sennet, A. (2016). Ambiguity. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2016/entries/ambiguity/

Sharma, N. (2008). Sensemaking handoff: When and how? *Proceedings of the American Society for Information Science and Technology*, *45*(1), 1–12. https://doi.org/10.1002/meet.2008.1450450234

Shipman, F. M., & Marshall, C. C. (1999). Formality Considered Harmful: Experiences, Emerging Themes, and Directions on the Use of Formal Representations in Interactive Systems. *Computer Supported Cooperative Work (CSCW)*, *8*(4), 333–352. https://doi.org/10.1023/A:1008716330212

Shneiderman, B. (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *Proceedings of the 1996 IEEE Symposium on Visual Languages*, 336.

Shneiderman, B. (2003). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In B. B. Bederson & B. Shneiderman (Eds.), *The Craft of Information Visualization* (pp. 364–371). Morgan Kaufmann. https://doi.org/10.1016/B978-155860915-0/50046-9

Shneiderman, B., & Plaisant, C. (2006). Strategies for evaluating information visualization tools: Multi-dimensional in-depth long-term case studies. *Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, 1–7. https://doi.org/10.1145/1168149.1168158

Shrinivasan, Y. B., & van Wijk, J. J. (2008). Supporting the analytical reasoning process in information visualization. *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*, 1237. https://doi.org/10.1145/1357054.1357247

Shrinivasan, Y. B., & Van Wijk, J. J. (2008). Supporting the Analytical Reasoning Process in Information Visualization. *Proceeding of the Twenty-Sixth Annual CHI*

Conference on Human Factors in Computing Systems, 1237. https://doi.org/10.1145/1357054.1357247

Shrinivasan, Y. B., & Van Wijk, J. V. (2009). Supporting Exploration Awareness in Information Visualization. *IEEE Computer Graphics and Applications*, *29*(5), 34–43. https://doi.org/10.1109/MCG.2009.87

Smith, P. J., Bennett, K. B., & Stone, R. B. (2006). Representation Aiding to Support Performance on Problem-Solving Tasks. *Reviews of Human Factors and Ergonomics*, *2*(1), 74–108. https://doi.org/10.1177/1557234X0600200104

Smith, P. J., & Hoffman, R. R. (Eds.). (2017). *Cognitive Systems Engineering: The Future for a Changing World* (1st ed.). CRC Press. https://doi.org/10.1201/9781315572529

Smithson, M. (1989). A Vocabulary of Ignorance. In M. Smithson (Ed.), *Ignorance and Uncertainty: Emerging Paradigms* (pp. 1–13). Springer. https://doi.org/10.1007/978-1-4612-3628-3_1

Smuc, M., Mayr, E., Lammarsch, T., Aigner, W., Miksch, S., & Gärtner, J. (2009). To Score or Not to Score? Tripling Insights for Participatory Design. *IEEE Computer Graphics and Applications*, *29*(3), 29–38. https://doi.org/10.1109/MCG.2009.53

Soares, A. G. M., Santos, C. G. R. dos, Mendonça, S., Carneiro, N. J. S., Miranda, B. P., Araújo, T. D. O. de, Freitas, A. A. de, Morais, J. M. de, & Meiguins, B. S. (2016). A Review of Ways and Strategies on How to Collaborate in Information Visualization Applications. *2016 20th International Conference Information Visualisation (IV)*, 81–87. https://doi.org/10.1109/IV.2016.69

Stapleton, G., Jamnik, M., & Shimojima, A. (2016). Effective Representation of Information: Generalizing Free Rides. In M. Jamnik, Y. Uesaka, & S. Elzer Schwartz (Eds.), *Diagrammatic Representation and Inference* (Vol. 9781, pp.

296–299). Springer International Publishing. http://link.springer.com/10.1007/978-3-319-42333-3_28

Statham, G., Campbell, S., & Klassen, K. (2012). The AvalX Public Avalanche Forecasting System. *International Snow Science Workshop*, 7.

Statham, G., Haegeli, P., Birkeland, K. W., Greene, E., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., & Kelly, J. (2010). The North American Public Avalanche Danger Scale. *International Snow Science Workshop*, 7.

Statham, G., Haegeli, P., Greene, E., Birkeland, K., Israelson, C., Tremper, B., Stethem, C., McMahon, B., White, B., & Kelly, J. (2018). A conceptual model of avalanche hazard. *Natural Hazards*, *90*(2), 663–691. https://doi.org/10.1007/s11069-017-3070-5

Statham, G., Holeczi, S., & Shandro, B. (2018, October 15). Consistency and Accuracy of Public Avalanche Forecasts in Western Canada. *International Snow Science Workshop*.

Sternberg, R. J., & Stemberg, K. (2012). *Cognitive Psychology (6th ed.)*. Cengage Learning.

Sterner, B. (2022). Explaining ambiguity in scientific language. *Synthese*, *200*(5), 354. https://doi.org/10.1007/s11229-022-03792-x

Sultana, S., Ahmed, S. I., & Rzeszotarski, J. M. (2023). Communicating Consequences: Visual Narratives, Abstraction, and Polysemy in Rural Bangladesh. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–19. https://doi.org/10.1145/3544548.3581149

Swaab, R. I., Postmes, T., Neijens, P., Kiers, M. H., & Dumay, A. C. M. (2002). Multiparty Negotiation Support: The Role of Visualization’s Influence on the Development of Shared Mental Models. *Journal of Management Information Systems*, *19*(1), 129–150. https://doi.org/10.1080/07421222.2002.11045708

Szafir, D. A., Haroz, S., Gleicher, M., & Franconeri, S. (2016). Four types of ensemble coding in data visualizations. *Journal of Vision*, *16*(5), 11–11. https://doi.org/10.1167/16.5.11

Techel, F., Mitterer, C., Ceaglio, E., Coléou, C., Morin, S., Rastelli, F., & Purves, R. S. (2018). Spatial consistency and bias in avalanche forecasts – a case study in the European Alps. *Natural Hazards and Earth System Sciences*, *18*(10), 2697–2716. https://doi.org/10.5194/nhess-18-2697-2018

Thomson, J., Hetzler, E., MacEachren, A., Gahegan, M., & Pavel, M. (2005). A typology for visualizing uncertainty. *Visualization and Data Analysis 2005*, *5669*, 146–158. https://doi.org/10.1117/12.587254

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, *185*(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vick, S. G. (2002). *Degrees of Belief—Subjective Probability and Engineering Judgement*. Renton.

Vuillemot, R., Boy, J., Tabard, A., Perin, C., & Fekete, J.-D. (2016). *LIVVIL: Logging Interactive Visualizations and Visualizing Interaction Logs* (p. 33). https://hal.inria.fr/hal-01535913

Wagner, M., Slijepcevic, D., Horsak, B., Rind, A., Zeppelzauer, M., & Aigner, W. (2019). KAVAGait: Knowledge-Assisted Visual Analytics for Clinical Gait Analysis. *IEEE Transactions on Visualization and Computer Graphics*, *25*(3), 1528–1542. https://doi.org/10.1109/TVCG.2017.2785271

Walny, J., Huron, S., Perin, C., Wun, T., Pusch, R., & Carpendale, S. (2018). Active Reading of Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 770–780. https://doi.org/10.1109/TVCG.2017.2745958

Ware, C. (2010). *Visual thinking for design*. Elsevier.

Ware, C. (2019). *Information visualization: Perception for design*. Morgan Kaufmann.

Ware, C. (2022). Chapter 6—Visual Objects, Words, and Meaning. In C. Ware (Ed.), *Visual Thinking for Information Design (Second Edition)* (pp. 107–124). Morgan Kaufmann.

Weick, K. E. (1995). *Sensemaking in organizations* (Vol. 3). Sage.

Willett, W., Heer, J., & Agrawala, M. (2007). Scented Widgets: Improving Navigation Cues with Embedded Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, *13*(6), 1129–1136. https://doi.org/10.1109/TVCG.2007.70589

Wood, J., Kachkaev, A., & Dykes, J. (2019). Design Exposition with Literate Visualization. *IEEE Transactions on Visualization and Computer Graphics*, *25*(1), 759–768. https://doi.org/10.1109/TVCG.2018.2864836

Woods, D. (2002). Escape from data overload. *Institute for Ergonomics. Ohio State University*.

Woods, D. D. (2002). Representations, Symbols, and Virtual Perceptual Fields. In *Escape from Data Overload*. Institute for Ergonomics. Ohio State University.

Wright, W., Schroh, D., Proulx, P., Skaburskis, A., & Cort, B. (2006). The Sandbox for Analysis: Concepts and Methods. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 801–810. https://doi.org/10.1145/1124772.1124890

Wu, A., Convertino, G., Ganoe, C., Carroll, J. M., & Zhang, X. (Luke). (2013). Supporting collaborative sense-making in emergency management through geovisualization. *International Journal of Human-Computer Studies*, *71*(1), 4–23. https://doi.org/10.1016/j.ijhcs.2012.07.007

Xiong, C., van Weelden, L., & Franconeri, S. (2019). The Curse of Knowledge in Visual Data Communication. *IEEE Transactions on Visualization and Computer Graphics*, 1–1. https://doi.org/10.1109/TVCG.2019.2917689

Xu, K., Attfield, S., Jankun-Kelly, T. J., Wheat, A., Nguyen, P. H., & Selvaraj, N. (2015). Analytic Provenance for Sensemaking: A Research Agenda. *IEEE Computer Graphics and Applications*, *35*(3), 56–64. https://doi.org/10.1109/MCG.2015.50

Xu, S., Bryan, C., Li, J. K., Zhao, J., & Ma, K.-L. (2018). Chart Constellations: Effective Chart Summarization for Collaborative and Multi-User Analyses. *Computer Graphics Forum*, *37*(3), 75–86. https://doi.org/10.1111/cgf.13402

Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—And is not—A desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition*, *41*(2), 229–241. https://doi.org/10.3758/s13421-012-0255-8

Yusoff, N. M., & Salim, S. S. (2015). A systematic review of shared visualisation to achieve common ground. *Journal of Visual Languages & Computing*, *28*, 83–99. https://doi.org/10.1016/j.jvlc.2014.12.003

Zhao, J., Glueck, M., Isenberg, P., Chevalier, F., & Khan, A. (2018). Supporting Handoff in Asynchronous Collaborative Sensemaking Using Knowledge-Transfer Graphs. *IEEE Transactions on Visualization and Computer Graphics*, *24*(1), 340–350. https://doi.org/10.1109/TVCG.2017.2745279

Zimmerman, J., & Forlizzi, J. (2014). Research Through Design in HCI. In J. S. Olson & W. A. Kellogg (Eds.), *Ways of Knowing in HCI* (pp. 167–189). Springer. https://doi.org/10.1007/978-1-4939-0378-8_8

Zuk, T., & Carpendale, S. (2007). Visualization of Uncertainty and Reasoning. In A. Butz, B. Fisher, A. Krüger, P. Olivier, & S. Owada (Eds.), *Smart Graphics* (pp. 164–177). Springer Berlin Heidelberg.

Appendix A.

This appendix includes supplementary materials for various studies. Each study with additional materials has its own section. Additional data files are detailed in each respective section.

Study 1

Table A.1. Table of example quotes and evidence gathered through observation in notes.

Theme	Sub-Theme	Definition	Part A Interview Quote	Part B Observed Evidence
Missing Info	Explicit	Missing information is explicitly represented in data.	"They [weather stations] get stuffed up by weather... rime or whatever... they just stop reporting." P4	
	Implicit	Missing information must be inferred from the situational context.	"When you're modeling the natural world, you take shortcuts and there's simplifications [...] the co-linearity between many of the avalanche problems […] they don't occupy fully independent places.... within our drive to communicate effectively, we sometimes have to have discussions about whether we want to be technically accurate, or whether we want to retain clarity [...] we sense that starts to get quite complicated. And so...frequently, we look for ways to simplify." P1	[Observed during discussions] "After having worked this job [Avalanche Canada] ... I sort of realize the big holes the operators leave in their writeups when it comes to work in this office using this information... because they are having face to face conversations... and maybe not putting that information into their writeup... saying this layer [of snow] does not exist in our area may not be helpful to them, but it really helps us here in this office." P8
Data Representativeness	Classification Overlap	Classifications are often not independent or mutually exclusive.	"When you're modeling the natural world, you take shortcuts and there's simplifications... The... major flaw of the conceptual model is the, the, the, the co-linearity between many of the avalanche problems and so... storms labs and wind slabs are heavily co-linear. So, you know, they, they don't occupy fully independent places." P1	[Observed during discussions] "Thinking about having all three [wind slab, storm slab, persistent slab], because they are so different right now.. There kind of always is a wind slab problem, but there was A LOT of wind so I want to capture that" P7
	Conservative Bias	Avalanche guides and professionals are conservative when faced with uncertainty in the field or in data.	"[If] we just like… puzzle it out, but then we still don't know. Like, I'll just start writing that… today, take a conservative approach" P2 "Oh, you know, this one operator was saying, you know, that they really found things touchy. So, I think I'm going to lean that way and be conservative, then there have been times where another forecaster would have said something like: 'Oh, well, you know, that... that person... Yeah,	[Observed unrecorded conversation about how to interpret an operators report considering their conservative bias caused by a recent incident involving clients] P3

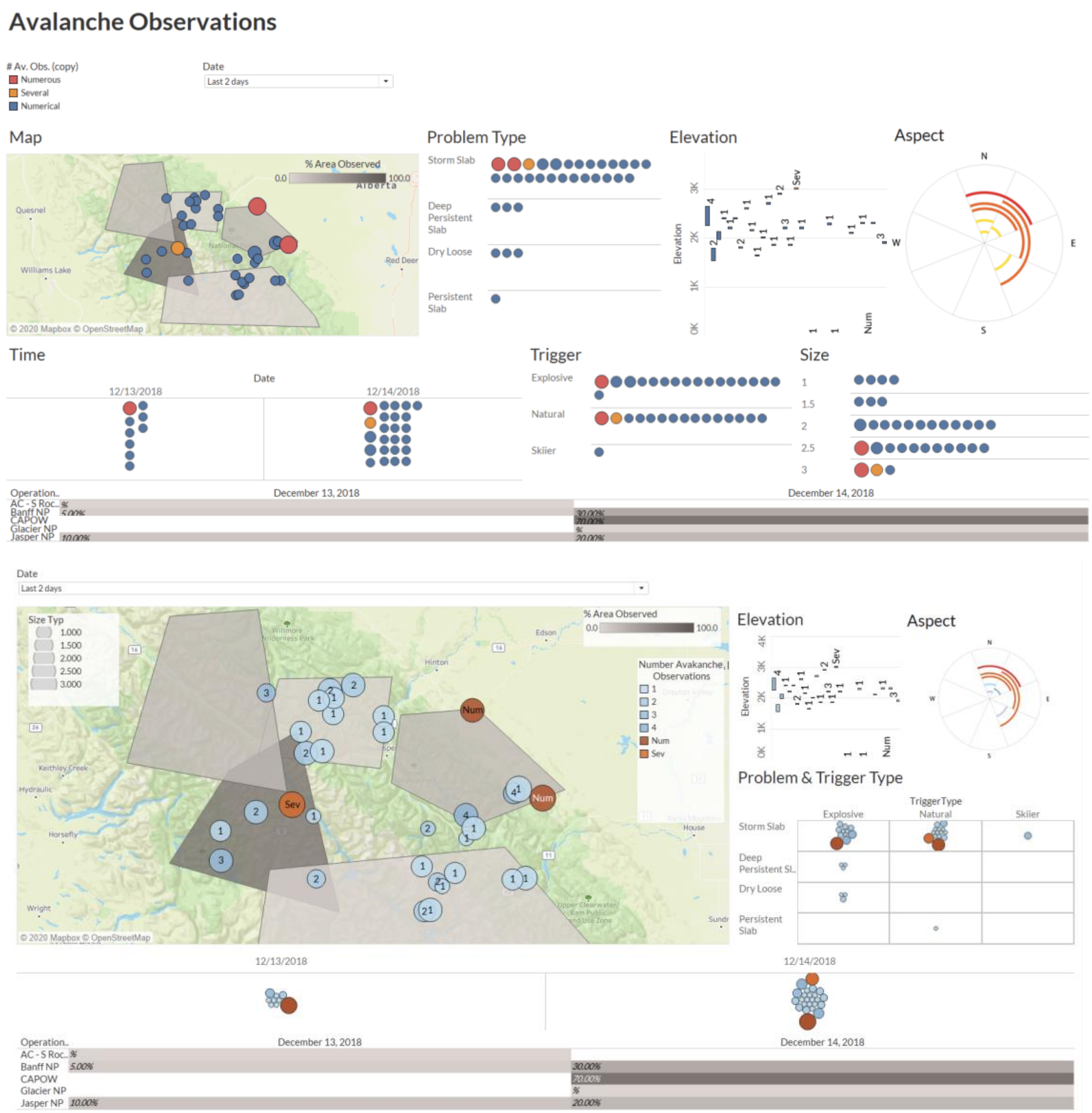
Theme	Sub-Theme	Definition	Part A Interview Quote	Part B Observed Evidence
			they always call that a little more than it actually is.' And then that may change… influence me to say: Okay, well, maybe I should not necessarily discredit it, but I put less weight into it." P3	
	Circumstantial Definitions	Official definitions and unofficial practices for reporting data depend on the situational context.	"The CAA courses do quite a good job of standardizing those kind of threshold amounts and stuff like that […but] people who have spent a lot of time on the coast, for example, may think 30 centimeters storm doesn't really do very much" P1	[Observed during discussions]"I like [Anonymized] point yesterday, wind slabs in the alpine are kind of like cornices that you find always... it is just a winter mountain hazard... it goes on the bulletin when it is elevated to more than normal caution..." P2
Analytic Practices	Subjective Hunches	Considering the behaviour, concerns, and hunches of others in the field to inform and guide analysis and interpretation.	"I might be reading that snowpack description like saying like… 'okay, are these guys still concerned about this?' That's what really matters to me more so than like the really nuanced low-level data." P2	[Debrief from video] "I feel good about who was in the operation. So, I felt that the test was valid and valid information that I should be thinking about." P3
	Immersion	Forecasters spend several days in forming a mental model through *undirected* review of contextual information.	"If I take over regions, I kind of try to ease into it. So, if I know that I've got five days on a certain amount of regions. I usually... The first day is just I don't... I may have questions in my mind, but I don't delve too deeply because it's just, it's too overwhelming." P4	[Debrief from video] "...It was just to give me an orientation to get my mental picture for forecasting in the Columbia [mountain region]. like where are we relative to the history... just a little bit of context... I don't know what that does for me exactly." P4
	Context-Seeking	*Directed* information search for supplementary contextual information.	"if I'm really struggling I'll like… look for keywords like "oh ya... skiing, like, steep terrain in the Alpine, up to 40 degrees and just exposed features. No problem." That tells me that not much is going on. Yeah, people are confident." P2	[Debrief from video] "...so I reviewed a few avalanches to understand what was driving those avalanches and why [Anonymized] added that persistent slab problem again." P6
	Mental Projection	Forecasters assimilate information by imagining and mentally visualizing the interactions of avalanche conditions, weather, terrain, and people.	"And, you know, by projecting yourself into the terrain. Actually, that's a technique that a lot of people use to help forecast and you know… kind of projecting yourself mentally, whether you close your eyes or you just have some kind of image of the kind of slopes, the kind of areas where the people are moving around in areas covered with trees, what the wind kind of might do, you know… I think it's pretty common to have some kind of, you know, visual representative little piece of terrain and you know, what people are doing. Users if you like, to help visualize, basically. So, you know, I think that experiential part, there is really relevant to the process." P1	
	Updating	Forecasters iterate over knowledge artifacts like their forecast as they conduct their analysis and	"So day one, if you... if you inherent forecasts you can kind of slightly tweak them all. And by day two, I'm usually grabbing one or two and doing like a pretty significant revamp of it. This is kind of where I put my voice in it	[Debrief from video] "I import yesterday's forecast... and I tweak my forecast, so it matches my nowcast" P6

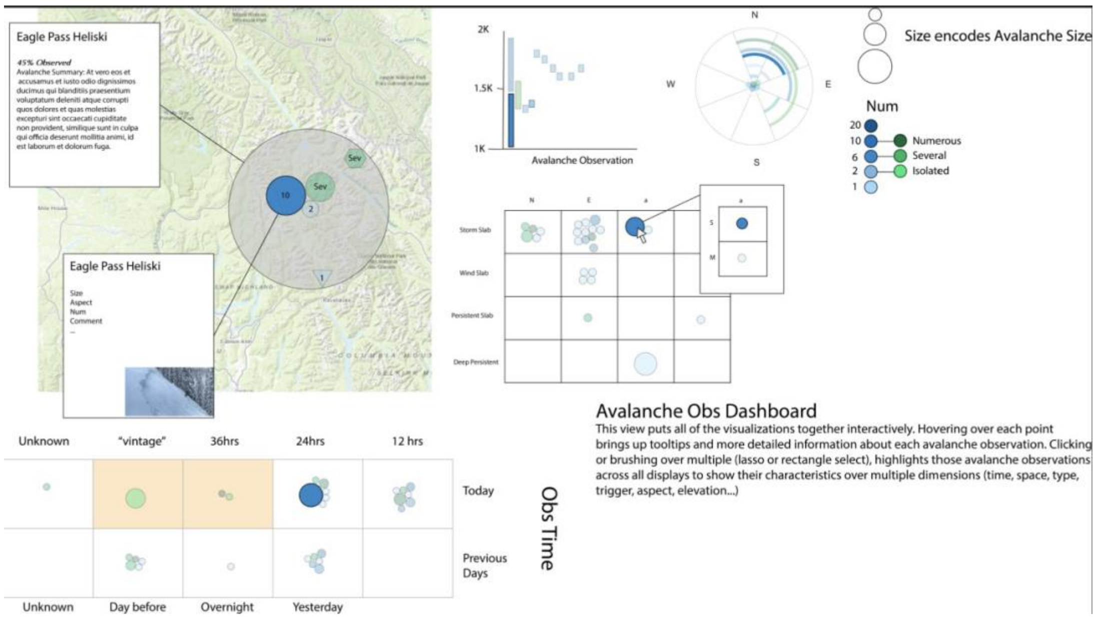
Theme	Sub-Theme	Definition	Part A Interview Quote	Part B Observed Evidence
		update their own mental models.	and any new insight that I might have in it. By day three, I probably… I've like updated all four of the forecasts. And I really put my voice into it. I have updated all the little tracking of the weak layers, that kind of thing. And then at that point, when you have a good feel for the regions like day four and five are a lot easier.” P2	
	Deliberate Omission	Forecasters manage information overload by ignoring certain data	“It's a funny balance because much of the individual data points are not very consequential. And so, from an efficiency point of view, it's better if you let your brain just merge those and average those out and have some kind of general assessment.” P1	[Debrief from video] “A result [snowpack test result] that’s a little alarming…that I wouldn’t expect under a low hazard… I am considering whether this layer is still a major concern. I decided my prior analysis still held [hazard is low]” P3
Analytic Challenges	Continuity	Forecasting relies on continuity of analysis and monitoring. Shift-changes disrupt this continuity.	“And there's a lot of variability in... in different people and what they... what sort of information and leave and... how much information they leave. But the idea is, you know that that's the first place I'll look and hoping that the forecaster the... previous forecaster has left enough information to start that picture, start getting an idea on what are the problems, and where are the uncertainties.” P3	[Observed during discussion] “I don’t think it was because it was a bad forecaster… it was because it had been in a few different hands... and then… someone did not have enough time to clarify it…” P2
	Translating Analysis	Forecasters struggle with communicating complex conditions with simple clarity to the public.	“You know, ‘snow turned into rain might might do this’, or it's… I try to… try to explain like the myriad of possibilities, there's just not that much room to explain things like that. And to put that kind of simply, that's a really challenging one” P2	[Observed during discussion] “It is only on 30 cms… to me it seems more reasonable to call it a persistent slab because it captures it on all aspects. “Wind slabs will be most reactive on the SH layer…” I find that confusing and I am thinking of ditching the wind slab problem” P7
	Lack of Good Representations	Forecaster lament a lack of good visual representations to alleviate some cognitive effort.	“Wind is another one of those things… a lot of the wind output is all in degrees. And so, it's like an hour and the value of the speed and then it's the degree… but you know, if you just read like 256 to 240, 220, 180… it's actually kind of hard… I actually keep this compass rose at my desk and look at it in my head to see where it is. But I think there's way better ways to visualize wind” P2	[Debrief from video] “I opened up 3 windows of InfoEx… one for today, one for yesterday and one for two days ago.” P6
	Lowering Danger Ratings	It is challenging for forecasters to lower danger ratings as data reveal instability rather than stability.	“it's because that ramp down is quite challenging. And, you know, it's... easy to go from considerable to high but it's hard to go down. It's easy to go from high back down to considerable but it's hard to go from considerable to moderate and it's even harder to go from moderate avalanche danger to low avalanche danger. Yeah, those	

Theme	Sub-Theme	Definition	Part A Interview Quote	Part B Observed Evidence
			are the hard steps… is actually coming down. And ...so anything to help us come to the conclusion that… you know, the problem is not acute anymore and we can ramp things down and come down off our you know, screaming to the high heavens that it's dangerous out there or whatever that is, I think is most helpful." P4	
Collaborative Sensemaking Strategies	Data Production	Forecasters facilitate collaborative work by producing hand-off notes and other internal knowledge artifacts.	"Or if I'm, if I'm uncertain about something... like that's what I might dive in more for the next forecaster [creating hand-off notes]... " It's like those little things like I am trying to take that ease and control that I have at day four or five, because I feel like I've got it under control and I give that to the next person so they don't feel like they have to do their process of discovery from ground zero essentially." P2	[Observation during discussions] "The notes [hand-off notes] thing is important… even if it's just a breadcrumb trail… even if it's just how you arrived at a conclusion… or even refer to {anonymized name} notes, it is just a carry-forward." P2
	Regular Discussions	Forecasters draw on each other's diverse knowledge through daily discussions.	"So, at two o'clock 230, we have our pow-wow where we all kind of go through our hazards and our problems. And, you know, does this make sense? Make sure that it's, it's, it's kind of like a peer review session."	[Observed and recorded several discussions during observations]
	Reaching out Directly	Forecasters call or email field operators for further information when faced with critical information gaps.	"...or am potentially missing something or I just don't feel comfortable with ...with what I have done, that's generally when I'll start picking the phone up and trying to find people in the area that can provide more… more insight." P3	[Observed forecaster making a phone call to an operator to inquire about conditions]
	Professional Exchange	Forecasters work with other agencies and operators to gain a deeper understanding of the nuances of how data are produced and what they mean.	"And the only way to really fully understand is to go and spend a bit of time with that operator. We try and facilitate that. We have professional exchanges go on. You know, we often go out. MOTI [The Ministry of Transportation and Industry] are really good partners with us. They're always happy for us to go and visit an operation." P1	

Study 2

Figure A.1. Intermediary low-fidelity prototypes of AvObs tool.





Study 3

In addition to the example surveymonkey form in the images below, the following spreadsheets containing data and supplemental information are included:

1. **Data_Dictionary.xlsx** definitions of fields and data attributes as well as participant metadata.
2. **Log_Data.xlsx** compiled video coding for sensemaking, interactions, and analytic actions.
3. **Assessments.xlsx** hazard assessments for both scenarios and survey feedback.

Figure A.2. Methodology diagram for study 3.

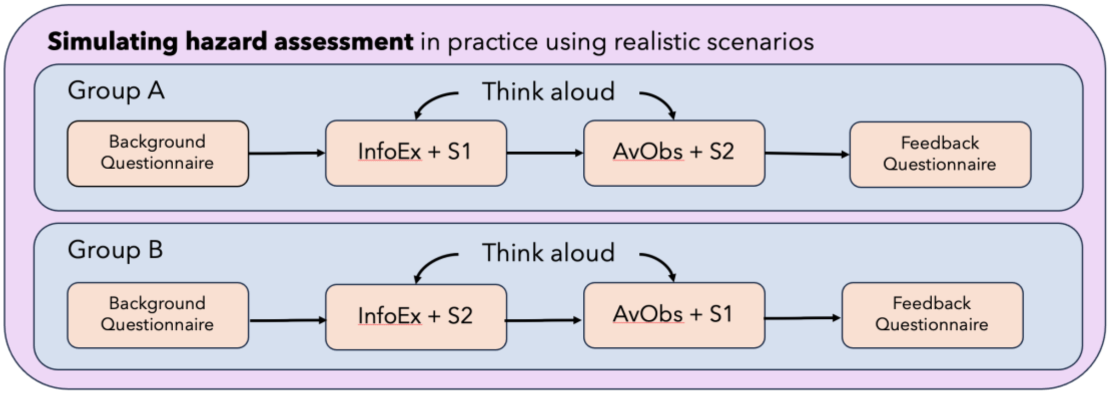
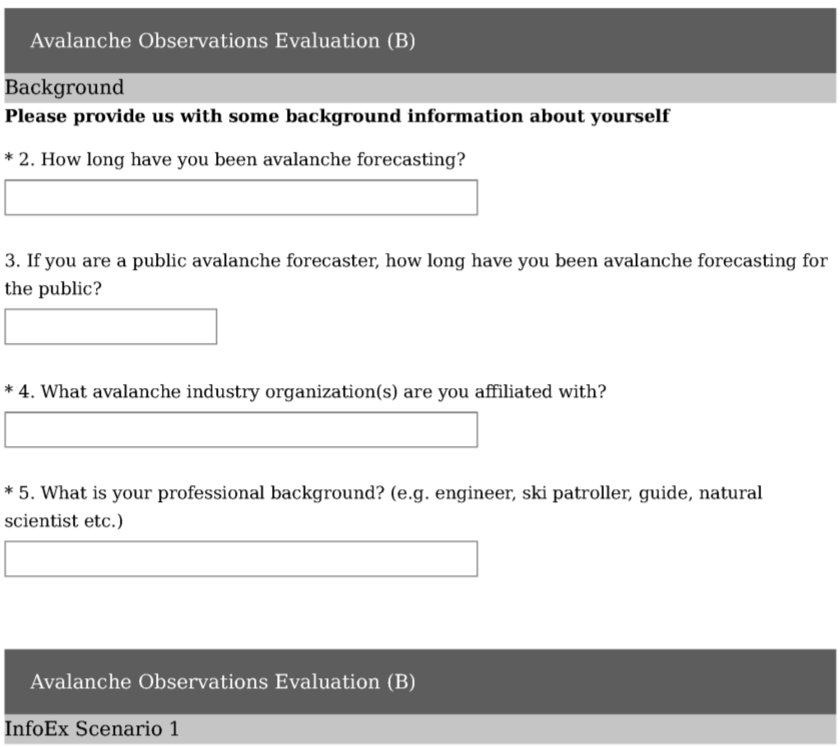


Figure A.3. The surveymonkey questionnaire presented to participants in Study 3 used (all following images)



Avalanche Observations Evaluation (B)

Background

Please provide us with some background information about yourself

* 2. How long have you been avalanche forecasting?

3. If you are a public avalanche forecaster, how long have you been avalanche forecasting for the public?

* 4. What avalanche industry organization(s) are you affiliated with?

* 5. What is your professional background? (e.g. engineer, ski patroller, guide, natural scientist etc.)

Avalanche Observations Evaluation (B)

InfoEx Scenario 1

Please start recording your screen now. If you don't have a screen recording tool available, please download ActivePresenter. Please ensure your recording software is functioning properly by capturing a short test video and talking into the computer microphone. Ensure that you are recording your entire workspace. This video recording is central to this study because it helps us understand how the tool supports the analytic process of avalanche hazard assessment.

You are tasked with completing two nowcasts for two scenarios. In each scenario you will receive the prior day's public bulletin along with overnight and daytime actuals. You will then analyze avalanche observations and avalanche summaries as if it were the end of the day and all InfoEx and MIN reports had been submitted already.

As you work through these scenarios, please speak into the microphone for the screen video recording and:

- verbalize your questions

"*What is this cluster here? Who is reporting here?*"

- verbalize your interpretations

"*This report is not representative of broader patterns and so I am not weighing it very heavily.*"

- verbalize your intentions and actions

"*I am going to look at some neighbouring regions as they have a fairly similar snow climate*"

- verbalize your reasoning for your actions

"*I hovered around those reports to get a sense of how the operators felt about the problem and what they saw*"

You may find that at times it is hard to talk at the same time as conducting your assessments. In these situations, try to retroactively explain what your thoughts and actions were. You may do so at natural pauses in your analysis.

In this first scenario, you will be using the InfoEx and the MIN to complete an AM avalanche hazard assessment (nowcast) for 07:00 on Tuesday March 10, 2020 in Glacier National Park. Glacier National Park produces bulletins in the morning for the day ahead. For the purposes of this study please do not go beyond the specified time-ranges and data sources.

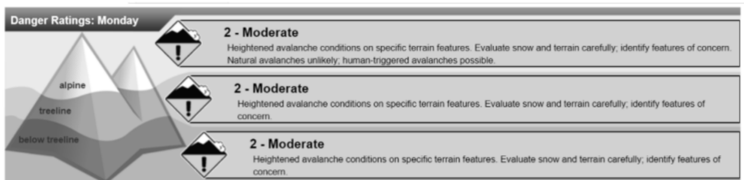
Please open the MIN & the InfoEx and use the "Avalanche Observations" and "Avalanche Summaries" tables only. You may view data between 2020-02-25 & 2020-03-09. You may view data from any region, but you are creating a nowcast for Glacier National Park.

Forecast issued Monday March 9th, 07:00 Valid through March 10th, 06:00 (Glacier National Park)

Happy Daylight Savings!

Give your minds extra time to assess slopes for the Feb 22nd weak layer.

We are into a period of lower probability/ high consequence!



Forecast	Tuesday	Wednesday
alpine	3 - Considerable	3 - Considerable
treeline	3 - Considerable	3 - Considerable
below treeline	3 - Considerable	2 - Moderate

Confidence: Moderate - Timing, track, or intensity of incoming weather system is uncertain

Problem 1: Persistent Slabs

Which Elevation?	Which Slopes?	Chance of Avalanches?	Expected Size?
			

The Feb 22 Surface Hoar/ sun crust layer is down ~90cm at TL in the W end of the park & 50-80cm in the east. It's a likely depth for skier triggering in shallower snowpack areas. This problem is decreasing in probability but remains high consequence

Travel and Terrain Advice

Be aware of the potential for wide propagations.
Avoid steep convexities or areas with a thin or variable snowpack.

Problem 2: Dry Loose

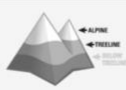
Which Elevation?	Which Slopes?	Chance of Avalanches?	Expected Size?
			

Unconsolidated new snow will be reactive to strong solar radiation, especially on steep, rocky features.

Travel and Terrain Advice

Minimize exposure to steep, sun exposed slopes when the solar radiation is strong.

Problem 3: Wind Slabs

Which Elevation?	Which Slopes?	Chance of Avalanches?	Expected Size?
			

Moderate to strong southerly winds this week have formed wind slabs in the alpine and exposed areas at treeline. If triggered, these wind slabs have the potential to step down to the Feb 22nd SH/ sun crust layer.

Travel and Terrain Advice

Use caution in lee and cross-loaded terrain near ridge crests.
If triggered the windslabs may step down to the Feb 22 SH layer with potential for large avalanches.

Weather Forecast

A cold front is moving southward across the province and will reach Rogers Pass by Tuesday morning.

Today: a mix of sun and cloud. Light west winds. Alpine high -11°C.

Tonight: cloudy with clear periods, no precipitation. Moderate SW winds. Alpine low -11°C

Tomorrow: 19cm of snow with moderate SW winds. Alpine high -8°C and freezing level at 900m.

Snowpack Discussion

25cm+/- of storm snow has been redistributed by moderate S'ly winds near ridgecrests in the alpine and exposed treeline. Below these slabs, the Feb 22nd persistent weak layer is now buried down 60-90cm, and consists of 3-7mm surface hoar on all aspects up to 2450m, and a crust on solar aspects. In some locations, this surface hoar sits on a crust.

Avalanche Activity Discussion

Numerous loose dry avalanches up to size 2.0 were observed yesterday on steep solar aspects, as well as several size 2.5 avalanches from the gullies of Mt. Macdonald.

On Friday, a snowmobile triggered size 2.5 slide occurred inside the E boundary of Glacier NP in the Bald Hills. Evidence indicated 2 sleds were involved and the group self extricated.

Weather Actuals as of 07:00 Tuesday, March 10th

Past 24 hour weather

Station	Maximum (°C)	Minimum (°C)	Snowfall (cm)	Snow Pack (cm)	Wind speed	Ridgetop wind direction
Fidelity 1905m	-5	-12	0	364	Moderate (26-40 km/h)	W
Rogers Pass 1315m	-1	-14	0	188	Moderate (26-40 km/h)	S

Complete an AM avalanche hazard assessment (nowcast) for 07:00 on Tuesday March 10, 2020

* 6. Avalanche Problem 1

- ◯ Storm Slab
- ◯ Wind Slab
- ◯ Persistent Slab
- ◯ Deep Persistent Slab
- ◯ Loose Dry Avalanche
- ◯ Loose Wet Avalanche
- ◯ Cornice

* 7. Problem 1: Aspect/Elevation Range (e.g. N-E TL-ALP)

8. Problem 1: Failure Plane

* 9. Problem 1: Spatial Distribution

* 10. Problem 1: Sensitivity to Triggers

* 11. Problem 1: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

* 12. Problem 1: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

13. Avalanche Problem 2

- ◯ Storm Slab
- ◯ Wind Slab
- ◯ Persistent Slab
- ◯ Deep Persistent Slab
- ◯ Loose Dry Avalanche
- ◯ Loose Wet Avalanche
- ◯ Cornice

14. Problem 2: Aspect/Elevation Range (e.g. N-E TL-ALP)

15. Problem 2: Failure Plane

16. Problem 2: Spatial Distribution

17. Problem 2: Sensitivity to Triggers

18. Problem 2: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

19. Problem 2: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

20. Avalanche Problem 3

- Storm Slab
- Wind Slab
- Persistent Slab
- Deep Persistent Slab
- Loose Dry Avalanche
- Loose Wet Avalanche
- Cornice

21. Problem 3: Aspect/Elevation Range (e.g. N-E TL-ALP)

22. Problem 3: Failure Plane

23. Problem 3: Spatial Distribution

24. Problem 3: Sensitivity to Triggers

25. Problem 3: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

26. Problem 3: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

* 27. Danger Rating

	Danger Rating
Alpine	
Treeline	
Below Treeline	

* 28. Headline

Avalanche Observations Evaluation (B)

Visualization Prototype Scenario 2

In the next scenario, you will be using the visualization prototype we have designed. Again, please do not go beyond the specified time-ranges and data sources. Make sure you are verbalizing your thoughts and actions for the screen recording.

Please open the AvID Sandbox and navigate to the AvObs tool. Preferably, please use the Chrome browser. Alternatively you may use Firefox, but all other browsers have not yet been tested with this prototype and we cannot guarantee it will function properly. You may view data between 2019-12-03 & 2019-12-17. You are forecasting for the North Columbias region. Avalanche Canada releases bulletins in the afternoon for the next day. In this scenario you will complete a PM avalanche hazard assessment (nowcast) for 18:00 on Tuesday, December 17th, 2019.

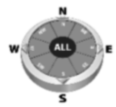
Forecast issued Monday Dec 16, 18:00 Valid through Tuesday Dec 17, 16:00 (North Columbia)

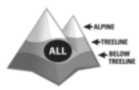
Snowfall combined with moderate to strong southwest wind Monday night through Tuesday will increase avalanche danger to HIGH.

Tuesday	
Alpine	4 - High
Treeline	3 - Considerable
Below Treeline	3 - Considerable
Wednesday	
Alpine	3 - Considerable
Treeline	3 - Considerable
Below Treeline	3 - Considerable
Thursday	
Alpine	3 - Considerable
Treeline	3 - Considerable
Below Treeline	3 - Considerable

Travel and Terrain Advice

- Minimize exposure during periods of heavy loading from new snow and wind.
- Storm slabs in motion may step down to deeper layers resulting in large avalanches.
- Be aware of the potential for large avalanches due to the presence of a persistent slab.

Avalanche Problem 1: Storm Slab			
What Elevation?	Which Slopes?	Chances of Avalanches?	Expected Size?
			
Snow accumulating above a buried layer of weak surface hoar will create a touchy storm slab problem. The conditions are primed for human triggered avalanches.			

Avalanche Problem 2: Persistent Slab			
What Elevation?	Which Slopes?	Chances of Avalanches?	Expected Size?
			
A couple of weak layers formed in late November and early December are now sitting about 1 m below the surface. This layer may present as surface hoar, a crust, facets or a combination of those, depending on elevation and aspect.			

Avalanche Summary

On Saturday there were 3 reports of skier triggered persistent and storm slab avalanches running on surface hoar buried 25-50 cm deep. These ranged from size 1 to 2 and were on north, west and south aspects between 1400 and 2000m. Two of these were remotely triggered (from a distance). There was also a report of a snowmobile triggered size 2 persistent slab avalanche on an east aspect at 2300m

Expect to see an increase in natural avalanche activity as storm snow accumulates through Monday night and Tuesday.

Snowpack Summary

25-40 cm of recent snow has accumulated above a widespread layer of large, feathery surface hoar crystals. New snow is expected to continue to accumulate above this layer over the next few days, making it more sensitive to human-triggering.

A weak layer formed in late November is now buried around 1 m plus below the surface. This is the layer of concern relating to the persistent slab avalanche problem. The weak layer may present as surface hoar, a crust, facets or a combination, depending on elevation and aspect. Below this, a variety of crusts from late October are buried deeper in the snowpack.

Weather Summary

MONDAY NIGHT: Snow, accumulation 10-15 cm, moderate to strong southwest wind, alpine high temperature -9 C.

TUESDAY: Flurries, accumulation 10-15 cm, moderate to strong southwest wind, alpine high temperature -4 C.

WEDNESDAY: Flurries, accumulation 5-10 cm, moderate south wind, alpine high temperature -6 C.

THURSDAY: Scattered flurries, moderate southwest winds, alpine high temperature -7 C.

Confidence: *Moderate*

- Uncertainty is due to the timing, track, & intensity of the incoming weather system.

Weather Actuals

Monday Night: 5 to 15 cm of snow fell with moderate to strong southwest wind. Alpine low temperature was -13 C.

Tuesday: Flurries throughout the day produced 5 to 15 cm of snow across the region with moderate to strong southwest wind. Daytime alpine high temperature was -5, and it's currently snowing.

Complete a PM avalanche hazard assessment (nowcast) for 18:00 on Tuesday December 17th, 2019

* 29. Avalanche Problem 1

- ◯ Storm Slab
- ◯ Wind Slab
- ◯ Persistent Slab
- ◯ Deep Persistent Slab
- ◯ Loose Dry Avalanche
- ◯ Loose Wet Avalanche
- ◯ Cornice

* 30. Problem 1: Aspect/Elevation Range (e.g. N-E TL-ALP)

31. Problem 1: Failure Plane

* 32. Problem 1: Spatial Distribution

* 33. Problem 1: Sensitivity to Triggers

* 34. Problem 1: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

* 35. Problem 1: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

36. Avalanche Problem 2

- ◯ Storm Slab
- ◯ Wind Slab
- ◯ Persistent Slab
- ◯ Deep Persistent Slab
- ◯ Loose Dry Avalanche
- ◯ Loose Wet Avalanche
- ◯ Cornice

37. Problem 2: Aspect/Elevation Range (e.g. N-E TL-ALP)

38. Problem 2: Failure Plane

39. Problem 2: Spatial Distribution

40. Problem 2: Sensitivity to Triggers

41. Problem 2: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

42. Problem 2: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

43. Avalanche Problem 3

- ◯ Storm Slab
- ◯ Wind Slab
- ◯ Persistent Slab
- ◯ Deep Persistent Slab
- ◯ Loose Dry Avalanche
- ◯ Loose Wet Avalanche
- ◯ Cornice

44. Problem 3: Aspect/Elevation Range (e.g. N-E TL-ALP)

45. Problem 3: Failure Plane

46. Problem 3: Spatial Distribution

47. Problem 3: Sensitivity to Triggers

48. Problem 3: Size

	Minimum Size	Typical Size	Maximum Size
Sizes			

49. Problem 3: Likelihood

	Minimum Likelihood	Typical Likelihood	Maximum Likelihood
Likelihoods			

* 50. Danger Rating

	Danger Rating
Alpine	
Treeline	
Below Treeline	

* 51. Headline

Avalanche Observations Evaluation (B)

Exit Survey

We would like to know your thoughts about how the prototype served your avalanche hazard assessment process.

52. Which features of the prototype did you find most useful and why?

53. Which features of the prototype did you find least useful and why?

54. Is there anything that the InfoEx system provides that you find is missing from the prototype? If so, what is it and why is it useful to you?

55. Are there any features or functionalities that you would like to see added to the prototype?

* 56. Describe how the way spatial data (tenures, locations, circle locations) was represented did or did not help in your assessment process. Is there anything you would change? If so, how?

* 57. Describe how the way temporal data (age ranges and direct observations) was represented did or did not help in your assessment process. Is there anything you would change? If so, how?

* 58. Describe how the way other data (# of avalanches, avalanche character, trigger types etc.) were represented did or did not help in your assessment process? Is there anything you would change? If so, how?

* 59. Please enter your initials here so that we can link your video file to this survey.

You may now stop your recording.

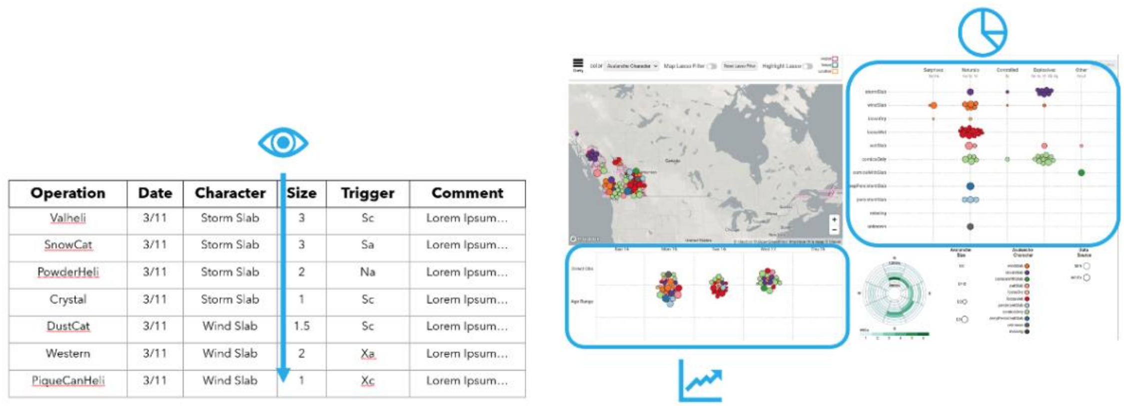
Please export your file to a video file. If you are using ActivePresenter, you may find instructions on how to do so here.

Please save the file with your initials and the date in the following format: SN_2021_07_21.mp4

It may take a few hour for the software to render to video so make sure you leave your computer on for this.

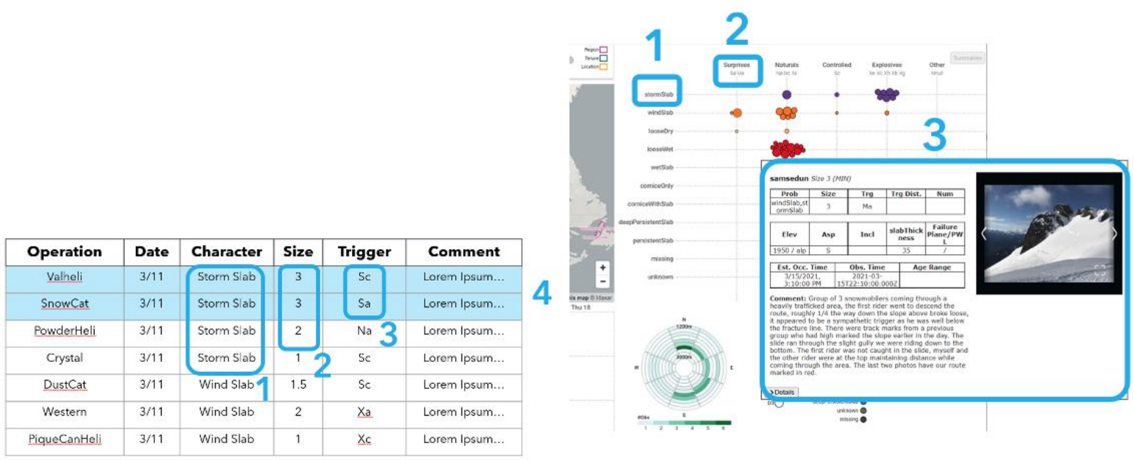
After the video has rendered, upload it to the following repository.

Figure A.4. Diagram of how analytic action of ‘pattern’ was identified in either toolset



Operation	Date	Character	Size	Trigger	Comment
Valheli	3/11	Storm Slab	3	Sc	Lorem Ipsum...
SnowCat	3/11	Storm Slab	3	Sa	Lorem Ipsum...
PowderHeli	3/11	Storm Slab	2	Na	Lorem Ipsum...
Crystal	3/11	Storm Slab	1	Sc	Lorem Ipsum...
DustCat	3/11	Wind Slab	1.5	Sc	Lorem Ipsum...
Western	3/11	Wind Slab	2	Xa	Lorem Ipsum...
PiqueCanHeli	3/11	Wind Slab	1	Xc	Lorem Ipsum...

Figure A.5. Diagram of how analytic action of ‘drill-down’ was identified in either toolset



Operation	Date	Character	Size	Trigger	Comment
Valheli	3/11	Storm Slab	3	Sc	Lorem Ipsum...
SnowCat	3/11	Storm Slab	3	Sa	Lorem Ipsum...
PowderHeli	3/11	Storm Slab	2	Na	Lorem Ipsum...
Crystal	3/11	Storm Slab	1	Sc	Lorem Ipsum...
DustCat	3/11	Wind Slab	1.5	Sc	Lorem Ipsum...
Western	3/11	Wind Slab	2	Xa	Lorem Ipsum...
PiqueCanHeli	3/11	Wind Slab	1	Xc	Lorem Ipsum...

Figure A.6. Diagram of how analytic action of ‘serial’ was identified in either toolset

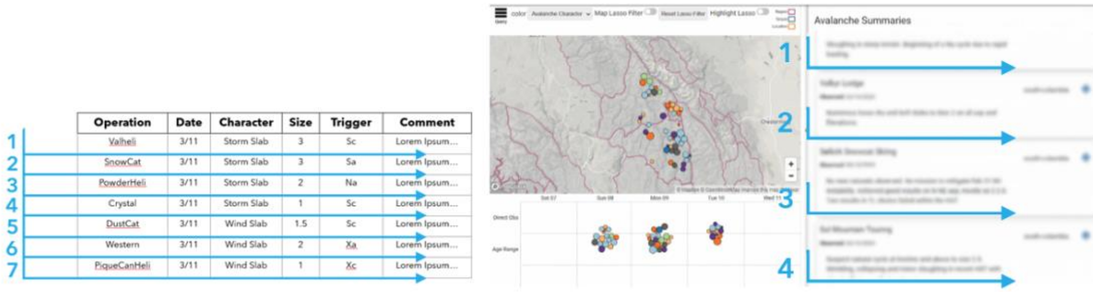


Figure A.7. Diagram of how analytic action of ‘focus’ was identified in either toolset

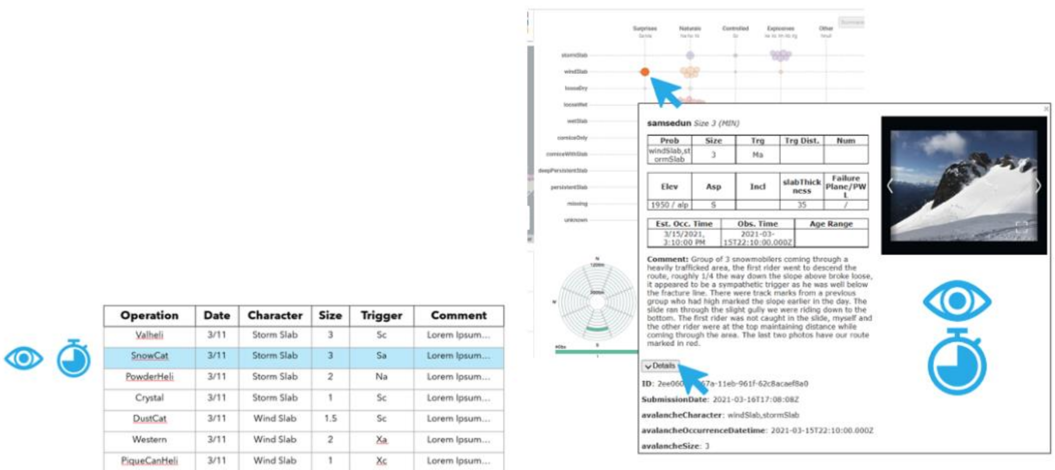


Figure A.8. Diagram of how analytic action of ‘gisting’ was identified in either toolset

Operation	Date	Character	Size	Trigger	Comment
Valheli	3/11	Storm Slab	3	Sc	Lorem Ipsum…
SnowCat	3/11	Storm Slab	3	Sa	Lorem Ipsum…
PowderHeli	3/11	Storm Slab	2	Na	Lorem Ipsum…
Crystal	3/11	Storm Slab	1	Sc	Lorem Ipsum…
DustCat	3/11	Wind Slab	1.5	Sc	Lorem Ipsum…
Western	3/11	Wind Slab	2	Xa	Lorem Ipsum…
PiqueCanHeli	3/11	Wind Slab	1	Xc	Lorem Ipsum…