I'm Not Sure: Designing for Ambiguity in Visual Analytics

Stan Nowak*
School of Interactive Arts
and Technology
Simon Fraser University

Lyn Bartram†
School of Interactive Arts
and Technology
Simon Fraser University

Abstract

Ambiguity, the state in which alternative interpretations are plausible or even desirable, is an inexorable part of complex sensemaking. Its challenges are compounded when analysis involves risk, is constrained, and needs to be shared with others. We report on several studies with avalanche forecasters that illuminated these challenges and how targeted visualization designs can better support ambiguity. Like many complex analysis domains, avalanche forecasting utilizes highly heterogeneous and incomplete data where the relevance and meaning of such data is context-sensitive, dependant on the knowledge and experiences of the observer, and is mediated by the complexities of communication and collaboration. In this paper, we characterize challenges of ambiguous interpretation emerging from *data*, *analytic process*, and *collaboration and communication*, and describe several management strategies for ambiguity. Drawing on our findings, we explore initial visual analytics design approaches to explicitly address ambiguity in complex sensemaking and risk analyses.

Index Terms: Human-centered computing—Visualization—Visualization theory, concepts and paradigms

1 Introduction

Our work addresses the challenges of complex and collaborative sensemaking in risk management: in particular, the domain of avalanche forecasting responsible for analysis and prediction of snow avalanches endangering human life and infrastructure. As is the case with explaining and predicting other hazards (such as weather or natural disasters), avalanche forecasting involves the consideration and evaluation of alternative potential explanations that account for data [22] and the communication of these predictions to audiences widely varying in expertise. In this way, sensemaking is deeply about managing ambiguity, the state of multiple alternative meanings, beyond simply accounting for missing information. Existing forecasting tools and procedures do not capture all the cognitive work forecasters do [41], motivating the design of better tools to support them.

Visual analytics, "the science of analytical reasoning facilitated by interactive visual interfaces" [12], is ideally suited to address the ambiguous sensemaking needs of avalanche forecasters. While visualization research has largely been devoted to quantified uncertainty or data uncertainty [7, 9, 13, 16, 25, 26, 37, 40, 53, 60, 62] researchers are now starting to explore broader issues of uncertainty related to reasoning [74] such as the interpretation of implicit errors [45, 55], the importance of "hunches" in data interpretation [35] or the role of alternatives in visual analysis [36]. We add to this growing body of work in our exploration of the challenges of sensemaking under ambiguity in risk analyses and the consequent implications for visualization designs.

*e-mail: Snowak@sfu.ca
†e-mail: Lyn@sfu.ca

The work reported in this paper focuses on two complementary threads of more effectively accommodating ambiguity in risk analysis and prediction. First, we seek a formative understanding of the challenges of ambiguity in complex and critical sensemaking. Through a set of studies with Avalanche Canada aimed at understanding the central challenges of avalanche forecasting we immediately discovered that ambiguity plays a critical role in sensemaking and at the same time presents challenges for both individual and collaborative analysis and communication. From our findings we characterize different sources of ambiguity and strategies to deal with it, grouped into issues of interpretation that deal with *data*, *analytic process*, and *collaboration and communication*.

Second, we use these findings to inform potential visual analytics design solutions as both interim tools and design probes to explore how potential visual analytics solutions can better serve the challenges of ambiguous interpretation involving heterogeneous data-generating processes. Deploying these solutions in practice as design probes serves ecological assessment that further surfaces the challenges and affordances of supporting ambiguity in reasoning about risk in the collaborative and critical environment of forecasting. Our design findings highlight both the effective potential of visualizations and the caveats. Key issues are the importance of multiple levels of data granularity, appropriate context, the need for analytic provenance, and the ability to capture both data and insights throughout the process.

The key takeaway of our research is that ambiguity is distinct from data uncertainty, requiring solutions that go beyond reduction or removal. It is an essential component of sensemaking, but at the same time presents specific challenges for analysis, collaboration, and communication. We argue that ambiguity can and should be designed **for** and not **away** [50], that even simple design choices can serve to support or impede sensemaking involving ambiguity, and that there is a need for more explicit ambiguity support in visual analytics tools. In this paper, we contribute:

- Insights from 3 qualitative studies with avalanche forecasters surfacing issues of ambiguity in sensemaking;
- A characterization of sources and strategies for ambiguity in risk analysis and sensemaking; and
- A preliminary exploration of visual analytics design approaches to address ambiguity.

2 Public Avalanche Forecasting

Public avalanche forecasters assess avalanche hazards and communicate the associated risks to the public through daily bulletins. These natural disasters endanger the safety of humans and infrastructure and require careful professional assessment to inform risk management in mountainous avalanche-prone areas. Forecasters try to predict how present or future instabilities within snowpack may react to natural triggers, such as the weight of new snow, or human triggers, such as the weight of a skier [44]

Avalanche forecasting is continuous and distributed across teams [41] of forecasters who monitor avalanche conditions over an entire winter season, iteratively updating their understanding with new

information [44]. While many forecasters have the benefit of working in the field and directly observing avalanche conditions, public avalanche forecasters work remotely and rely heavily on field reports produced by other organizations [50]. In Canada, such reports are shared in the Canadian Avalanche Association's Industry Information Exchange (InfoEx) [19] by avalanche safety 'operators', such as those overseeing railway or transportation corridors, ski resorts, and helicopter skiing operations among others. While these data are structured and defined using formal measurement and reporting guidelines [2], they are gathered using a targeted sampling rather than a random sampling approach [44]. Operators actively seek instabilities in the snow. Consequently, forecasters have to glean enough context about this process to understand what such data mean (who reported it, what happened, etc.).

Another challenge stems from the sparsity of data. For example, remote weather stations used to validate meteorological forecasts [50] are very sparsely distributed when compared to the variability and heterogeneity of mountain weather [38]. Forecasters mentally simulate the interactions of mountainous terrain and weathers systems and their effects on snowpack from limited data. This imaginative and speculative ability is a mark of competence and expertise in avalanche forecasting [1] as well as weather forecasting [57].

Forecasters formalize their judgements of avalanche hazards using a variety of qualitative measures such as a danger scale, likelihood scale, potential destructive size, as well as different avalanche problem types [65]. These assessments are then communicated to the public through daily bulletins that are supplemented with additional risk communications such as advice about how to avoid avalanche hazards. The public varies in levels of expertise and consequently varies in how they interpret even simple elements of bulletins such as danger scales [17, 64]. Public avalanche forecasters rely heavily on their knowledge, experience, and expert judgment to assess and communicate avalanche hazards. The challenges of complexity, varied interpretation, and uncertainty are similar to those involved in risk prediction and communication of other extreme weather events and natural disasters [6].

3 Related Work

3.1 Sensemaking and Ambiguity

Sensemaking — the process by which meaning is constructed based on available information and experience — is precipitated by information or events that violate expectations or are uncertain and ambiguous [42, 70]. It is characterized by complexity. Complexity involves dynamically evolving rules and interacting parts [21] where comprehensive understanding is intractable [28] due to the epistemological limitations of human observation [18]. These limitations mean that complexity is more effectively dealt with holistically rather than through mechanistic reduction to the sum of parts. Sensemaking addresses complexity and the concomitant uncertainties through the flexible construction of narratives [14] where informational *cues* help determine what is relevant and which narratives or explanations are coherent or acceptable to consider [10].

This "narrative mode" of thinking describes how signs, symbols, representations, and their relationships are tied together into coherent personal narratives authored by the observer [4]. A novel is merely ink until it is read by someone and the same applies to the analysis of data. Subplots and micro-narratives involving prior knowledge and personal experiences are involved in the reading and making sense of visualizations [51]. Just as a story involves competing narratives, so too does sensemaking in general. This is because sensemaking generally starts with an existing explanation that is challenged by a viable alternative [30]. Sensemaking is thus more about resolving multiple potential meanings (ambiguity), rather than just accounting for missing or uncertain information.

3.2 Sensemaking and Risk Prediction

Risk management work faces real-world time constraints, ill-defined goals, distributed tasks and responsibilities, uncertainty, and decision-making demands. The engineering of technological solutions to deal with these issues requires close consideration of the cognitive processes involved [23, 63]. Frequently in these domains, for example in weather forecasting [22], several targeted sensemaking strategies are employed. Generally, these involve the setting of expectations to direct attention to cues that can signal threats and a concurrent sensitivity to cues that deviate from these expectations [71].

One example relevant to the forecasting of avalanches is *anticipatory thinking*: a functional form of mental preparation for potential risks including those that may be highly unlikely but could result in severe consequences [33]. Attention is actively managed and directed to subtle and context-sensitive cues that may signal threats. There are several types of anticipatory thinking. One, problem detection, describes the process by which observers first become aware of an issue that may require a course of action [31, 32]. The ability to detect problems depends on how rich an observer's understanding of relevant patterns to compare against data is. This "pattern matching" often involves monitoring multiple patterns or "frames" concurrently. Anticipatory thinking also involves "trajectory tracking", the extrapolation of trends into multiple alternative future scenarios as well as planning for them. The imagination, exploration, and planning for alternative scenarios is also known as mental simulation [29]. These processes are vulnerable to psychological factors or biases such as a tendency to explain away disconfirming evidence. However, studies with expert weather forecasters show such biases are countered by through the active adoption of a skeptical stance in analysis [31].

3.3 Ambiguity in Visualization Research

Visualization research has a longstanding tradition of characterizing uncertainties relevant to the design of visual analytics systems [8, 39, 40, 68, 74]. Most visualization research has focused on data uncertainties, but many acknowledge the importance and role of interpretation and knowledge in uncertainty [16, 27, 39, 54, 74]. MacEachren discusses ambiguity through the lens of organizational decision-making describing it as a "*lack of an appropriate 'frame of' reference through which to interpreted the information*" and describes equivocality as stemming from the diversity of possible interpretations [39]. Meanwhile, Boukhelifa et al. define ambiguity in terms of multiplicities in the relationship between entities and names in data as well as the differences in interpretation between collaborators [8]. Liu et al. present a framework for the exploration, interpretation, and management of alternatives in visual analytics [36]. They group alternatives into three types: cognitive (e.g. hypotheses, mental models, and interpretations), artifact (e.g. data, models, representations, or tools), and execution (e.g. methods, code, and parameters). Ambiguity is most closely related to their concept of cognitive alternatives. Most prominent in existing visualization research is the discussion of ambiguity in collaborative visual analytics where sharing of analysis is often incomplete, lacking context, and therefore ambiguous [20].

There is much more to analysis than what is explicit in data. Data are incomplete records of the phenomena they are intended to represent and require prior knowledge as well speculation. This is closely related to the notion of "implicit errors", which are errors inherent to a dataset but not explicitly represented within it [45, 55]. To better support sensemaking around implicit errors associated with infectious disease statistics, McCurdy et al. used structured annotations to help expert clinicians externalize knowledge about these errors [45]. In an application for archeological analyses, Panagiotidou et al. developed visualization tools that explicitly represented implicit errors [55]. Lin et al. use the term *data hunch* to describe "a person's knowledge about how representative data is of a phe-

nomenon of interest" and how issues like credibility, inclusion and exclusion criteria, or directionality and magnitude of biases are considered in the analysis of data [35]. The authors outline a design space for externalizing data hunches.

4 Approach

We conducted three studies (Figure 1) with Avalanche Canada, a Canadian public avalanche forecasting organization, to understand the central challenges of avalanche forecasting and how visual analytics solutions could address them. We began with semi-structured interviews to understand how forecasters perceive and describe the challenges of their work (Study 1). We then conducted field observations of forecasters on site. Concurrently, we video-recorded forecaster's workstations and debriefed them about analytical reasoning involving the use of existing technologies and tools using a method called Cued-Recall and Debrief [52] (Study 2). This set of observations corroborated and enriched our understanding of the themes we identified in the interview study (Figure 2). We then designed two visualization prototypes in collaboration with the avalanche forecasters and conducted retrospective interviews using these prototypes as design probes (study 3). The purpose of this last study was to better understand hoq visual analytics interventions can address the challenges of ambiguity.

We frame our findings according to issues of ambiguity dealing with data, analytic process, or collaboration and communication throughout this paper. Data are incomplete records of the phenomena they represent and in turn require nuanced and varying interpretations depending on the needs and goals of analysis. Considering and evaluating alternative interpretations is an essential part of sensemaking. The analytic process of judging and adopting alternative interpretations presents potential analytic paths through data. These paths can be difficult to navigate as much of analysis is not explicitly captured. Finally, collaborators each hold unique perspectives and thus alternative interpretations, while communication simplifies complexity, removes context, and produces ambiguity. We apply this structure to illustrate how evidence and findings from our studies informed our characterization of ambiguity and to enrich our discussion of the design implications for potential visual analytics solutions.

Studies 1 and 2 were conducted on-premises at Avalanche Canada while study 3 was conducted remotely. In total, 12 avalanche forecasters participated in our studies (P1-P5 participated in Study 1, P2-P8 in Study 2, and P2-P6 / P9-P12 in Study 3). 10 were male and 2 were female reflecting the gender balance of the organization and industry. The forecasters came from varied and mixed backgrounds. 8 had a background in professional mountain guiding, 3 in engineering, 2 in natural sciences, and 2 in business and communications.

5 Study 1: Examining Forecaster Work Challenges

5.1 Procedure

We conducted semi-structured interviews with 5 professional avalanche forecasters on Avalanche Canada premises in Revelstoke, British Columbia. We 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. Participants were asked questions like: "*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?*". The interviews were audio-recorded and then transcribed.

5.2 Analysis

Transcripts were concurrently segmented 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. Inter-rater reliability was measured with one other coder using a sample of the transcript. Simple agreement for high-level themes was .89, Cohen's Kappa was .81, and Krippendorff's Alpha was .82. For the sub-themes, simple agreement was .75, Cohen's Kappa was .70, and Krippendorff's Alpha was .71.

5.3 findings

5.3.1 Data

Forecasters informed us that one of the key challenges they face is 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... because they, you know... They get stuffed up by weather... rime or whatever... they just stop 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

Similarly, forecasters have to use 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 adopted through extensive experience and knowledge. "*The CAA 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, for example, may think a 30 centimeter storm doesn't really do very much*" P1 Meanwhile, as is common in 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 undue complexity in 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*[...] *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

The nuances of evidential reasoning and interpretation of data in avalanche forecasting also extend to the risk-based **conservative bias** that avalanche professionals have. Some professionals may be more or less conservative, and forecasters have to factor in such considerations when weighing evidence. "[A]*nother forecaster would have said something like: '*[...]*they always call that a little more than 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

5.3.2 Analytic Process

Forecasters utilize a variety of analytic processes that draw on their prior knowledge, experiences as well as contextual clues and involve imagination and speculation. They told us they use **mental projection**, a process of imagining oneself in the field to understand conditions and their risk implications. "*...that'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 that station.* [...] *I'm like taking an input and then adjusting it for myself.*" P2 It might also involve

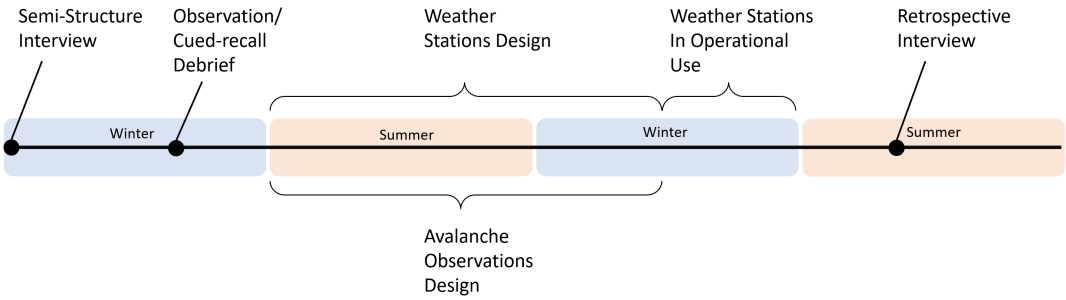

Figure 1: A timeline of studies.

simulating alternative future scenarios and their risk implications. "*If things are a little bit unusual, I do 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 reported having to spend time **immersing** 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. "*...I would suggest, at least 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. "*If I'm really struggling, 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 Forecasters describe forecasting as bayesian-like because they are constantly updating their mental models with new information and **deliberately omitting** weak or redundant evidence. As they conduct their assessments, they iteratively **update** knowledge artifacts like the public bulletin to match their current understanding. "*I'm pretty iteratively making the 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 Unlike forecasters, operators directly observe avalanche conditions in the field and thus have a richer understanding of the complexities involved. As a result, forecasters use subtle cues that can reveal the **subjective hunches** of operators and help forecasters appropriately frame their understanding of avalanche conditions. "*I might be reading that snowpack description, 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

5.3.3 Collaboration and Communication

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. "[M]*y 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 230, 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. "[W]*hether 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. "[T]*here'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 This is seen as a separate and additional task often completed at the end of the day when forecasters are fatigued. This is why artifacts used in support of hand-off and collaboration are 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 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 like 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

6 Study 2: Observing Avalanche Analytics

The purpose of study 2 was to observe forecaster's workplace behaviours and their use of technology. We expected to find a richer understanding of the challenges faced by forecasters and how visual analytics interventions may help.

6.1 Procedure

We conducted field observations on Avalanche Canada premises for a week. Field notes, as well as audio recordings of daily discussion, were gathered. At the same time, we gathered observations using cued-recall debrief (CRD), a situated recall method developed

	S1 Theme	S1 Sub-Theme	Definition	S2 Observed Evidence (O = Observation, C = CRD)
Data	Missing Info	Explicit	Missing information is explicitly represented in data.	
		Implicit	Missing information must be inferred from the situational context.	O
	Data Representativeness	Classification Overlap	Classifications are often not independent or mutually exclusive.	O
		Conservative Bias	Avalanche professionals are conservative when faced with uncertainty in the field or in data.	O
		Circumstantial Definitions	Official definitions and unofficial practices for reporting data depend on the situational context.	O
Analytic Process	Analytic Practices	Subjective Hunches	Considering the behaviour, concerns, and hunches of others in the field to inform and guide analysis and interpretation.	C
		Immersion	Forecasters spend several days forming a mental model through *undirected* review of contextual information.	C
		Context-Seeking	*Directed* information search for supplementary contextual information.	C
		Mental Projection	Forecasters assimilate information by imagining and mentally visualizing the interactions of avalanche conditions, weather, terrain, and people.	
		Updating	Forecasters iterate over knowledge artifacts like their forecast as they conduct their analysis and update their own mental models.	C
		Deliberate Omission	Forecasters manage information overload by ignoring certain data	C
	Analytic Challenges	Lack of Good Representations	Forecasters lament a lack of good visual representations to alleviate some cognitive effort.	C
		Lowering Danger Ratings	It is challenging for forecasters to lower danger ratings as data reveal instability rather than stability.	
Collaboration and Communication		Continuity	Forecasting relies on the continuity of analysis and monitoring. Shift-changes disrupt this continuity.	O
		Translating Analysis	Forecasters struggle with communicating complex conditions with simple clarity to the public.	O
	Collaborative Sensemaking Strategies	Data Production	Forecasters facilitate collaborative work by producing hand-off notes and other internal knowledge artifacts.	O
		Regular Discussions	Forecasters draw on each other's diverse knowledge through daily discussions.	O
		Reaching out Directly	Forecasters call or email field operators for further information when faced with critical information gaps.	O
		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.	

Figure 2: Table with thematic codes developed in study 1 (semi-structured interviews) and applied to study 2(field observations and cued-recall debrief). Inter-rater reliability was measured in Study 1. Simple agreement for high-level themes was .89, Cohen's Kappa was .81, and Krippendorff's Alpha was .82. For the sub-themes, simple agreement was .75, Cohen's Kappa was .70, and Krippendorff's Alpha was .71. Our thematic codes are organized and color coded according to their relevance to different sources of ambiguity.

for use in complex decision-making contexts [52] and adapted for human-computer interaction [5]. 7 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. At regular intervals, video recordings were reviewed to identify timestamps where forecasters exhibited behaviours relevant to our research interests. At the end of the workday, recordings were played back to forecasters at marked timestamps, and forecasters were asked to explain their thought processes and actions. We asked questions like: "*Can you explain what you were doing and thinking here?*" These debrief interviews were video recorded and transcribed.

6.2 Analysis

We applied the thematic coding scheme developed in study 1 to notes and transcripts in study 2. This allowed us to compare what forecasters say and what they do in order to enrich our understanding. Thematic coding was applied by one coder in two passes.

6.3 Findings

We found that forecasters rely heavily on text tables and information from disparate web-based sources. We observed forecasters repurposing web-browser features to accomplish simple analytic tasks. For instance, one forecaster opened several days of data in successive windows to investigate temporal patterns and make comparisons. This suggested forecasters could benefit from dedicated analytic tools to support such tasks. To our surprise, we found that the visualizations present in existing systems were seldom used. While it was clear the forecasters could benefit from dedicated analytic tools, the overwhelming use of text tables indicated this representational form held some comparative advantage in sensemaking.

6.3.1 Data

Several of the daily **discussions** we observed dealt with the topic of how to interpret particular reports. 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, 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

We note how much of organizational knowledge relevant to the interpretation of data is oral tradition. It is exchanged during the practice of work when applied to the particularities and complexities of situations faced. A whole host of procedural and tacit knowledge is exchanged through the shared practice and environment of work.

6.3.2 Analytic Process

When debriefing forecasters about their workday we found they relied more on **subjective hunches** of operators that they personally trusted and were more familiar with. This factored into how evidence was weighed. "*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

We 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 [61]. "*...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 Many aspects of analysis, particularly those dealing with personal knowledge, trust, or early parts of sensemaking, may be difficult to articulate and consequently share with others.

6.3.3 Collaboration and Communication

Forecasters don't just iterate over their own bulletin over the course of the day, they carry forward the previous day's bulletin even if another forecaster wrote it. We 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 [anonymized] *added that persistent slab problem again.*" P6

Forecasters are sensitive to changes made to the public bulletin and actively try to understand the reasons behind the changes. The bulletin provides an analytic scaffolding coordinating the forecasters' collaborative sensemaking processes. This suggests collaboration may be improved using tools that make evidential reasoning processes more explicit.

7 Study 3: Exploring Visualization Solutions

Informed by our general understanding of the forecasters' work practices, we set out to develop visualization prototypes to support core analytic forecasting tasks. These visualizations serve as design probes to further inform our understanding of how visual analytics interventions may serve to address the challenges of ambiguity in this domain. The first prototype aggregates meteorological observations from remote weather stations in order to help forecasters validate the previous day's weather forecast as well as to monitor evolving weather systems in real-time. The second prototype uses field-reported avalanche observations produced by avalanche safety operations sharing data in the InfoEx. Avalanche observations are treated as key indicators of avalanche hazards in avalanche forecasting.

7.1 Procedure

We designed and developed both prototypes through several iterations from paper sketches to computational implementation in collaboration with avalanche forecasters at Avalanche Canada. Both tools were evaluated using think-aloud protocol throughout the design process, exploring how tools support reasoning. The tools were also evaluated using retrospective interviews, soliciting forecasters' reflections about how the tools served avalanche forecasting work. The avalanche observations prototype used simulated synthetic and historical data from past seasons and was never used operationally. The weather stations prototype used real-time data and was used operationally in the second half of the winter forecasting season. 7 forecasters had input on the design and development of tools while one simply commented on their experiences using them. At the end of the forecasting season, we conducted semi-structured interviews asking forecasters to reflect on the prototypes, how they addressed the challenges of data, how they affected their work, and what needs remain unfulfilled. Interviews were conducted remotely using video conference tools. We used our prototypes as artifacts in the interview to prompt and probe the forecaster's reflections. The interviews were video-recorded and transcribed. We summarize our key findings with quotes extracted from transcripts below

7.2 Findings

7.2.1 Weather Stations

Analysis of weather station telemetry presents issues of **data uncertainty**. They are sparsely distributed relative to the large spatial areas they are used to represent [38] and they are subject to a variety of sensor and transmission errors caused by environmental factors. There is no comprehensive automated quality assurance procedure that accounts for all possible errors in the data [47]. Diagnosing 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. Our visualization prototypes employed visualizations of aggregate measures, multiple granularities of data, interactions including brushing and filtering, and tooltips to view the details of individual weather station telemetry (Figure 3). Our visualization prototype introduced a **new and unfamiliar analytic 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

For many, who are used to seeing **raw data** in a **tabular format** this raised issues of **trust** as they could not apply the same visual scanning strategies to detect errors in data. "[I]*t largely stems from the trustworthiness of the data* [... addressing the use of spreadsheets] *I like things in their raw format just for my own sake, my own just my own stamp of approval.* [...] *I guess it's easy for my eyes to decode differences or irregularities. You should be able to visualize the data and get the same output. I don't know why. I just use tables.*" P1 Others also used raw data tables but did so to **scaffold the learning** of data processing mechanics and the affordances of the visualizations as analytic tools. "[...] *having that* [raw data table] *side by side with the visualization helped me to interpret: Okay, what's the visualization trying to tell me here?*" P4

Forecasters who adopted the visualizations more readily and found they provided them with a richer and deeper understanding of meteorological phenomena than traditional data tables. Drawing a historical comparison to the role of computers in meteorology, forecasters view visualizations as a stepping stone in a **transitional phase** towards more data-driven modeling. "[T]*here 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

Our findings suggest that showing granular data is important not only to address the ambiguity of diagnosing multiple potential sources of error but also the ambiguity of adopting new analytic approaches. Hasty transitions risk the loss of a host of implicit procedural knowledge that may not be supported by new approaches. This can cause issues of trust. Showing raw data alongside more abstracted views of data can aid comprehension of new tools and allow users to evaluate their affordances.

7.2.2 Avalanche Observations

Our second prototype (Figure 4) used daily field-reported avalanche observations shared by avalanche safety operators on the InfoEx platform. These tables are generally viewed in a tabular format. Again, we attempted a conventional visualization approach using simple numerical aggregations like counts. This was of little use to the forecasters, however. We found several issues requiring disaggregated views of data. First, **how data standards are used and interpreted depends on the situational context.** "*...the InfoEx system and the standards... they kind of define the box that we all work in* [...] *how you use them... context drives that. You might use a certain approach... data that are obviously within that general framework or box that we've created, but you might not use them exactly the same way.*" P12. This meant that such data have ambiguous expressions, where the same data value may correspond to multiple meanings depending on context and the communicative intent of the report author. In addition, data are gathered using a targeted sampling approach rather than a random sampling approach. The data generating process is not uniform across the dataset and as a result, this challenges the methodological soundness of aggregate measures. "*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

Forecasters wanted to see individual reports while at the same time being able to discern general patterns in the data. To address this design constraint we used glyphs with circle marks representing individual reports in a packed layout within a variety of visualization contexts. Circle marks were encoded with core structured data attributes of reports. The size of circles encoded typical avalanche

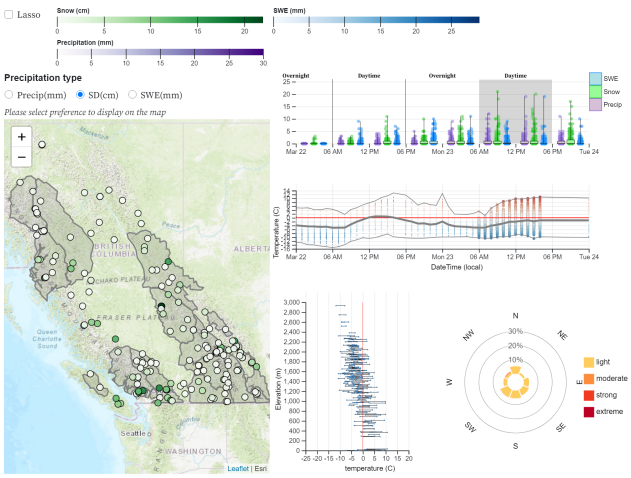

Figure 3: Prototype used to analyze weather station telemetry. Sensor and transmission errors present multiple sources of data uncertainty at individual weather stations that must be discerned through human interpretation. While tooltip interactions reveal the necessary details to support sensemaking about this form of ambiguity, the use of numerical aggregates to provide an overview presented a new and unfamiliar method to the forecasters. This raised issues of trust, but the use of tables showing raw data aided translating prior analytic processes into this new representation.

size and the color encoded the number of observed avalanches. Two color maps were used to distinguish numerical and categorical values reflecting the need to preserve raw forms of data. Brushing and linking as well as tooltip interactions reveal contextual details allowing forecasters to discern how to interpret individual reports. This glyph-based approach operates at multiple scales of resolutions allowing forecasters to visually aggregate data to discern patterns. Glyphs are known to support several visual aggregation operations such as summarizing data, detecting outliers, detecting trends, or segmenting data into clusters [67].

Throughout the design process, forecasters were concerned about visualizations giving them a **false sense of precision** disarming the level of scrutiny forecasters usually apply to these data. To address this concern, we chose to use perceptually weak visual encodings like color and size rather than position. We conjecture this lack of precision introduces visual complexity that induces additional effort, attention, and careful consideration of data. This beneficial form of difficulty has been discussed in the context of geovisualization and risk-based decisions [11] and is well-documented in studies of human learning [73].

Forecasters reported feeling satisfied with how the designs represented and supported their analytical processes. "[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

The properties of these data, like their ambiguous expressions or the varying data-generating processes, challenge conventional visualization approaches. However, a targeted design approach that reveals contextual details that cue nuanced interpretations can better support the necessary scrutinizing sensemaking processes such data demand. Our design may not be the only or even the optimal way to represent these data, but it serves to highlight that ambiguity can and should be designed for and that even simple visualization designs can better support the challenges of ambiguous sensemaking.

8 Discussion

Throughout our 3 studies, we found that critical issues of ambiguity arise in three contexts: the data, the process of analysis, and the challenges of communicating both data and interpretation to both co-workers and the general public. We unpack the role of ambiguity, the concomitant challenges, and strategies used to deal with ambiguity in each of these contexts. Our findings highlight the need for more effective design interventions. We discuss each in turn.

8.1 Data

Ambiguity emerges from data because they are incomplete simplifications of the complex phenomena they represent. Ambiguity may be involved in the expression of data or how representative data are of phenomena of interest. Whether reasoning about multiple types of data uncertainty in weather station telemetry or what field-reported avalanche observations mean for avalanche conditions more broadly, forecasters use their knowledge, experience, and cues within the data to explore plausible explanations that account for what they see. **Provoking** alternative interpretations serves a productive purpose in analysis.

Forecasters try to **capture** relevant nuances of interpretation about specific data through daily discussions. Often this might serve to **disambiguate** meaning by providing an optimal or appropriate framing for the data. For instance, the understanding that weather stations at windy locations will need adjustment when trying to understand broader wind patterns. We note that the forecasters' corpus of organizational knowledge is predominantly oral tradition exchanged in application to the immediate demands of work. Such a mechanism for knowledge exchange is vulnerable to information loss.

8.2 Analytic Process

Ambiguity serves a productive purpose in analytic processes but at the same time poses challenges for the management and navigation of analyses. Alternative interpretations are explored as part of sensemaking often taking the form of alternative scenarios in risk analysis and risk prediction. Either through mental visualization or explicit sketches, forecasters **provoke** and imagine alternative scenarios to explore potential risks or coherent explanations of data.

Additionally, the judgments and analytic choices made during analysis represent 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, the evidential reasoning

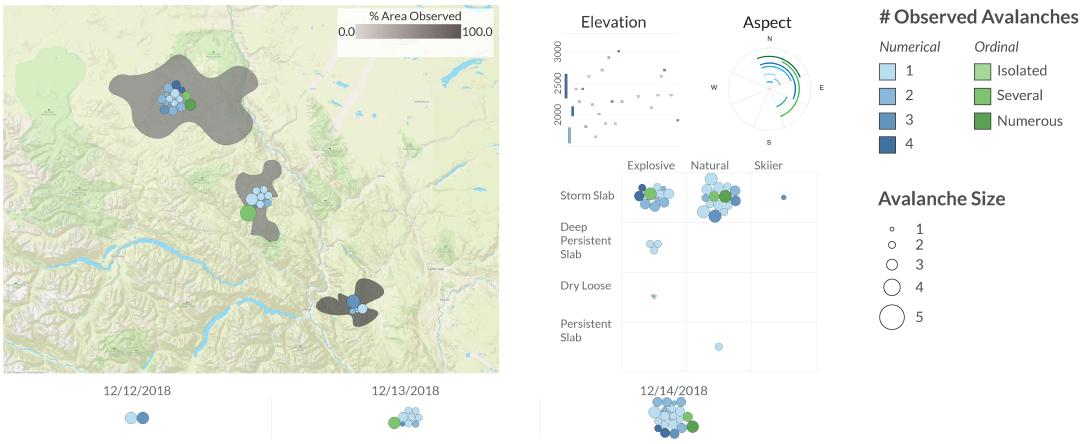

Figure 4: Prototype used to investigate field-reported avalanche observations. Glyphs with circles in a packed layout are placed in a variety of visualization contexts. Brushing and linking across multiple views and tooltip interactions reveal contextual details necessary for nuanced interpretation of individual reports. Perceptually weak visual encodings are intended to avoid a false sense of precision and encourage scrutiny of data.

process behind their judgments is often left **uncaptured** and may be difficult to reconstruct. This is likely to hide what work is already completed and what work remains to be done.

8.3 Collaboration and Communication

Collaborators each provide a unique interpretive lens, while communicating analysis simplifies complexity and removes context. To better support collaborative sensemaking forecasters use strategies like regular discussions or hand-off notes to exchange knowledge and **disambiguate** how to interpret each other's assessments. Forecasters try to **capture** their own reasoning processes and interpretations. However, given the additional effort of this task, and difficulty in anticipating what may be relevant, such information is often not completed leaving forecasters to speculate about their colleagues' reasoning processes.

Forecasters translate their own complex understanding of avalanche conditions in simple terms to ensure that members of the public, whether novice or expert, can apply appropriate risk-management strategies. In doing so, forecasters **mitigate** the effects of potential scenarios the public might encounter or the confusion that might result from overly technical communications. Quite often, this means **reconciling** alternatives. For instance, in a situation where two avalanche problem types require the same risk mitigation strategies, forecasters will use one of them and supplement any further guidance that might be necessary using plain and actionable language. The myriad of ways to communicate hazards presents its own form of ambiguity. Moreover, individual forecasters differ in how they judge avalanche hazards and apply assessments [34, 66].

8.4 Design Implications

Through our design study we learned that visual displays of such heterogeneous **data** should reveal the contextual details necessary to discern appropriate interpretations. Abstractions like numerical aggregations can occlude details which can impede sensemaking. Instead, we recommend using designs such as unit visualizations that capitalize on perceptually-enabled visual aggregations or those showing the relevant granularity of data alongside numerical aggregations. This allows alternative interpretations to be **provoked** when trying to understand how data come to represent a phenomenon of interest.

We argue that design solutions need to extend beyond the representation of existing data. The ambiguities of data interpretation may be aided by more **explicit ambiguity capture mechanisms**. These may for instance serve to characterize ambiguities through the externalization of relevant knowledge. We must take care these interventions remain lightweight and contextually anchored to avoid undue effort. We draw inspiration from the concept of "active reading", where knowledge generated during the process of reading is captured with external representations such as computationally-enabled markup and annotations [46]. Researchers have demonstrated that such techniques can be extended to the reading of visualizations [69] well as exploratory visual analysis [58]. Annotations are a common solution for dealing with ambiguity [8] and have also been employed to address implicit errors [45], suggesting annotations could be extended to more specifically address ambiguity. Other forms of markup [3] including annotations employed for the nuanced interpretation of data are often embedded in the ubiquitous spreadsheet, perhaps the most widespread analytic tool. Tables are flexible and allow direct interaction with data which might explain why users often turn to them to support complex sensemaking. The affordances of tables are well-suited to deal with the challenges of ambiguity and may serve to guide the design of visual analytics systems.

Such capture mechanisms present some practical challenges when scaling. Annotations tend to produce large amounts of unstructured data that can be difficult to reuse. Such data require a **schematization** mechanism to make them tractable for future reuse. Elicitation mechanisms may be structured ahead of time, for example through survey-like questionnaires. Meta-data gathered at the time of elicitation, such as timestamps or application states [43], may be exploited to give some structure. Alternatively, natural language processing mechanisms such as ontology-learning may lend themselves to schematizing such meta-data. However, we stress that the use of such algorithms should maintain transparency and give supervisory control to users. As we have learned in our design study, even simple statistical abstractions can obfuscate details paramount to reasoning about ambiguity. Further, highly complex technological solutions are more vulnerable to failure [71]. Consequently, the use of automation or algorithms should be carefully designed to make data processing transparent in support of human comprehension.

Ambiguity is often the start of a sensemaking process. At such early stages, understanding may be inchoate and difficult to articulate, calling into question the utility of highly detailed capture mechanisms such as annotations. In collaborative analysis, it is difficult to anticipate the needs of others. Collaborators might only form an intuition about a problem that may be important for others to be

aware of [31]. This is because the relevance of any such problem is context-sensitive [71]. Standardized protocols for sharing analysis often fail because designers of such protocols cannot adequately account for and predict all the unique information or complexity that might arise [56]. These considerations are important whether collaboration is with others or oneself at a future point in time.

There may be more simple capture mechanisms that can address the difficulties of articulating complexity. Passive capture mechanisms such as interaction logs provide one lightweight and context-sensitive solution. Interaction logs have been used to infer reasoning processes [15] and are frequently discussed as approaches for documenting analytic provenance [72]. Interaction logs, however, only show behaviours and are indirect indicators of reasoning processes. User-controlled markup may still be necessary to capture what is relevant. Researchers in clinical healthcare settings have supplemented handoff protocols with vague metrics like gut feelings about a patient, time spent with a patient, or how equipment in the room has been moved around to take advantage of practitioners' shared work culture and environment [49]. We can take inspiration from this work. To capitalize on the shared digital working environment, simple markup such as tagging of data or representations may be all that is necessary to signify ambiguity. Tags may signify important pieces of evidence, how evidence is weighed and relates to assessments, or may simply serve to raise awareness of ambiguity and prevent it from being lost potentially resulting in misinterpretation. Forecasters can use their shared working environment to maintain context and capture ambiguities without having to precisely articulate them. Awareness of uncertainty is critical to ensuring trust in findings [59]. We argue the same applies to awareness of ambiguity.

The use of explicit capture mechanisms raises the question of how long captured data should **persist** as part of the working environment. Such markup may only be relevant for one working session and one individual. Alternatively, It might be relevant across several working days and for multiple collaborators. Or, it might take a more permanent form in a corpus of organizational knowledge. The design of capture mechanisms should consider ways to control or account for the persistence of captured data.

The process of managing and exploring alternatives may also benefit from more explicit support. For example, exploring alternative risk scenarios related to instabilities in snow may be aided through the use of physical simulations. Physical simulations of snowpack structure are starting to be used in operational avalanche forecasting for hazard assessment [24, 48] and could be extended to allow user-driven exploration of how the snowpack would respond to a variety of triggers.

9 Limitations

We note that while our first study had additional coders to test reliability, data from subsequent studies was analyzed by one coder only. Our comprehension of the challenges that forecasters face was incorporated in prototypes within our design study and the feedback forecasters provided throughout our close collaboration served as a form of validation of our understanding. This presents obvious limitations in the reliability of our findings. However, such challenges are common in the development of long-term, qualitative, and ethnographically inspired research aimed at deep domain understanding.

10 Conclusion

We have presented findings from a set of qualitative studies with public avalanche forecasters. Our research highlights that ambiguity presents challenges and unmet needs in critical and complex complex sensemaking. We propose a formative characterization of ambiguity across three levels of abstraction in analysis: *data, analytic process,* and *collaboration and communication*. The key lesson of our research is that ambiguity should be explicitly considered and designed for. While even simple visualization design choices can serve to enable or impeded sensemaking around ambiguity, we argue for more targeted and explicit approaches. Our findings may inform research and the design of tools in other complex risk-management domains such as extreme weather forecasting or the forecasting of other natural disasters. This work represents a preliminary attempt to characterize ambiguity and define a design space for visual analytics, but many questions remain unexplored. Further study is necessary to evaluate our existing and proposed design solutions to more rigorously understand their impact and how they address the challenges of ambiguity.

Acknowledgments

The authors wish to thank A, B, and C. This work was supported in part by a grant from XYZ.

References

[1] L. Adams. A systems approach to human factors and expert decision-making within the canadian avalanche phenomena. 284. Publisher: Citeseer.

[2] C. A. Association and others. *Observation guidelines and recording standards for weather, snowpack and avalanches*. [Revelstoke, BC]: Canadian Avalanche Assoc.

[3] L. Bartram, M. Correll, and M. Tory. Untidy data: The unreasonable effectiveness of tables.

[4] L. R. Beach. Narrative thinking and decision making: How the stories we tell ourselves shape our decisions, and vice versa.

[5] T. Bentley, L. Johnston, and K. von Baggo. Evaluation using cued-recall debrief to elicit information about a user's affective experiences. In *Proceedings of the 17th Australia Conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future*, pp. 1–10.

[6] K. J. Beven, S. Almeida, W. P. Aspinall, P. D. Bates, S. Blazkova, E. Borgomeo, K. Goda, J. W. Hall, J. C. Phillips, M. Simpson, P. J. Smith, D. B. Stephenson, T. Wagener, M. Watson, and K. L. Wilkins. Epistemic uncertainties and natural hazard risk assessment. 1. a review of different natural hazard areas. doi: 10.5194/nhess-2017-250

[7] A. M. Bisantz, D. Cao, M. Jenkins, P. R. Pennathur, M. Farry, E. Roth, S. S. Potter, and J. Pfautz. Comparing uncertainty visualizations for a dynamic decision-making task. 5(3):277–293. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[8] N. Boukhelifa, M.-E. Perrin, S. Huron, and J. Eagan. How data workers cope with uncertainty: A task characterisation study. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pp. 3645–3656.

[9] K. Brodlie, R. A. Osorio, and A. Lopes. A review of uncertainty in data visualization. pp. 81–109. Publisher: Springer.

[10] A. D. Brown, P. Stacey, and J. Nandhakumar. Making sense of sense-making narratives. 61(8):1035–1062. Publisher: Sage Publications Sage UK: London, England.

[11] L. Cheong, C. Kinkeldey, I. Burfurd, S. Bleisch, and M. Duckham. Evaluating the impact of visualization of risk upon emergency route-planning. 34(5):1022–1050. Publisher: Taylor & Francis.

[12] K. A. Cook and J. J. Thomas. Illuminating the path: The research and development agenda for visual analytics.

[13] M. Correll and M. Gleicher. Error bars considered harmful: Exploring alternate encodings for mean and error. 20(12):2142–2151. Publisher: IEEE.

[14] G. Currie and A. D. Brown. A narratological approach to understanding processes of organizing in a UK hospital. 56(5):563–586. Publisher: Sage Publications.

[15] W. Dou, D. H. Jeong, F. Stukes, W. Ribarsky, H. R. Lipford, and R. Chang. Recovering reasoning processes from user interactions. 29(3):52–61. Publisher: IEEE.

[16] M. Fernandes, L. Walls, S. Munson, J. Hullman, and M. Kay. Uncertainty displays using quantile dotplots or cdfs improve transit decision-making. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–12.

[17] H. Finn. Examining risk literacy in a complex decision-making environment: A study of public avalanche bulletins.

[18] F. Grabowski and D. Strzalka. Simple, complicated and complex systems—the brief introduction. In *2008 Conference On Human System Interactions*, pp. 570–573. IEEE.

[19] P. Haegeli, J. Obad, B. Harrison, B. Murray, J. Engblom, and J. Neufeld. InfoexTM 3.0–advancing the data analysis capabilities of canada's diverse avalanche community. In *Proceedings of the International Snow Science Workshop*, vol. 29, pp. 910–917.

[20] J. Heer and M. Agrawala. Design considerations for collaborative visual analytics. 7(1):49–62. Publisher: SAGE Publications Sage UK: London, England.

[21] B. Hilligoss and S. D. Moffatt-Bruce. The limits of checklists: handoff and narrative thinking. 23(7):528–533. Publisher: BMJ Publishing Group Ltd.

[22] R. R. Hoffman, D. S. LaDue, H. M. Mogil, P. J. Roebber, and J. G. Trafton. *Minding the weather: How expert forecasters think*. MIT Press.

[23] E. Hollnagel and D. D. Woods. *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC press.

[24] S. Horton, S. Nowak, and P. Haegeli. Enhancing the operational value of snowpack models with visualization design principles. 20(6):1557–1572. Publisher: Copernicus GmbH.

[25] J. Hullman, X. Qiao, M. Correll, A. Kale, and M. Kay. In pursuit of error: A survey of uncertainty visualization evaluation. 25(1):903–913. Publisher: IEEE.

[26] J. Hullman, P. Resnick, and E. Adar. Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering. 10(11):e0142444. Publisher: Public Library of Science San Francisco, CA USA.

[27] A. Kale, M. Kay, and J. Hullman. Decision-making under uncertainty in research synthesis: Designing for the garden of forking paths. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–14.

[28] D. Kirsh. Thinking with external representations. 25(4):441–454. Publisher: Springer.

[29] G. Klein and B. W. Crandall. The role of mental simulation in problem solving and decision making. In *Local applications of the ecological approach to human-machine systems*, pp. 324–358. CRC Press.

[30] G. Klein, J. K. Phillips, E. L. Rall, and D. A. Peluso. A data–frame theory of sensemaking. In *Expertise out of context*, pp. 118–160. Psychology Press.

[31] G. Klein, R. Pliske, B. Crandall, and D. D. Woods. Problem detection. 7(1):14–28. Publisher: Springer.

[32] G. Klein, R. M. Pliske, B. Crandall, and D. Woods. Features of problem detection. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 43, pp. 133–137. SAGE Publications Sage CA: Los Angeles, CA. Issue: 3.

[33] G. Klein, D. Snowden, and C. L. Pin. Anticipatory thinking. In *Informed by Knowledge*, pp. 249–260. Psychology Press.

[34] B. Lazar, S. Trautmann, M. Cooperstein, E. Greene, and K. Birkeland. North american avalanche danger scale: Do backcountry forecasters apply it consistently. In *Proceedings ISSW*, pp. 457–465. Citeseer.

[35] H. Lin, D. Akbaba, M. Meyer, and A. Lex. Data hunches: Incorporating personal knowledge into visualizations.

[36] J. Liu, N. Boukhelifa, and J. R. Eagan. Understanding the role of alternatives in data analysis practices. 26(1):66–76. Publisher: IEEE.

[37] L. Liu, A. P. Boone, I. T. Ruginski, L. Padilla, M. Hegarty, S. H. Creem-Regehr, W. B. Thompson, C. Yuksel, and D. H. House. Uncertainty visualization by representative sampling from prediction ensembles. 23(9):2165–2178. Publisher: IEEE.

[38] J. Lundquist, M. Hughes, E. Gutmann, and S. Kapnick. Our skill in modeling mountain rain and snow is bypassing the skill of our observational networks. 100(12):2473–2490.

[39] A. M. MacEachren. Visual analytics and uncertainty: Its not about the data. Publisher: The Eurographics Association.

[40] A. M. MacEachren, A. Robinson, S. Hopper, S. Gardner, R. Murray, M. Gahegan, and E. Hetzler. Visualizing geospatial information uncertainty: What we know and what we need to know. 32(3):139–160. Publisher: Taylor & Francis.

[41] L. Maguire and J. Percival. Sensemaking in the snow: Exploring the cognitive work in avalanche forecasting. In *International Snow Science Workshop, Innsbruck, Austria*.

[42] S. Maitlis and M. Christianson. Sensemaking in organizations: Taking stock and moving forward. 8(1):57–125. Publisher: Routledge.

[43] A. Mathisen, T. Horak, C. N. Klokmose, K. Grønbæk, and N. Elmqvist. InsideInsights: Integrating data-driven reporting in collaborative visual analytics. In *Computer Graphics Forum*, vol. 38, pp. 649–661. Wiley Online Library. Issue: 3.

[44] D. McClung. The elements of applied avalanche forecasting, part i: The human issues. 26(2):111–129. Publisher: Springer.

[45] N. McCurdy, J. Gerdes, and M. Meyer. A framework for externalizing implicit error using visualization. 25(1):925–935. Publisher: IEEE.

[46] H. Mehta, A. Bradley, M. Hancock, and C. Collins. Metatation: Annotation as implicit interaction to bridge close and distant reading. 24(5):1–41. Publisher: ACM New York, NY, USA.

[47] E. Mekis, N. Donaldson, J. Reid, A. Zucconi, J. Hoover, Q. Li, R. Nitu, and S. Melo. An overview of surface-based precipitation observations at environment and climate change canada. 56(2):71–95. Publisher: Taylor & Francis.

[48] S. Morin, S. Horton, F. Techel, M. Bavay, C. Coléou, C. Fierz, A. Gobiet, P. Hagenmuller, M. Lafaysse, M. Ližar, and others. Application of physical snowpack models in support of operational avalanche hazard forecasting: A status report on current implementations and prospects for the future. 170:102910. Publisher: Elsevier.

[49] F. Mueller, S. Kethers, L. Alem, and R. Wilkinson. From the certainty of information transfer to the ambiguity of intuition. In *Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments*, pp. 63–70.

[50] S. Nowak, L. Bartram, and P. Haegeli. Designing for ambiguity: Visual analytics in avalanche forecasting. In *2020 IEEE Visualization Conference (VIS)*, pp. 81–85. IEEE.

[51] S. Nowak, L. Bartram, and T. Schiphorst. A micro-phenomenological lens for evaluating narrative visualization. In *2018 IEEE Evaluation and Beyond-Methodological Approaches for Visualization (BELIV)*, pp. 11–18. IEEE.

[52] M. M. Omodei and J. McLennan. Studying complex decision making in natural settings: using a head-mounted video camera to study competitive orienteering. 79(3):1411–1425. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[53] L. Padilla, M. Kay, and J. Hullman. Uncertainty visualization. Publisher: PsyArXiv.

[54] L. M. Padilla, M. Powell, M. Kay, and J. Hullman. Uncertain about uncertainty: How qualitative expressions of forecaster confidence impact decision-making with uncertainty visualizations. p. 3747. Publisher: Frontiers.

[55] G. Panagiotidou, R. Vandam, J. Poblome, and A. Vande Moere. Implicit error, uncertainty and confidence in visualization: an archaeological case study. Publisher: Institute of Electrical and Electronics Engineers.

[56] E. Patterson. Structuring flexibility: the potential good, bad and ugly in standardisation of handovers. 17(1):4–5. Publisher: BMJ Publishing Group Ltd.

[57] R. M. Pliske, B. Crandall, and G. Klein. Competence in weather forecasting. 40:68. Publisher: Cambridge University Press Cambridge, UK.

[58] H. Romat, N. Henry Riche, K. Hinckley, B. Lee, C. Appert, E. Pietriga, and C. Collins. ActiveInk: (th) inking with data. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.

[59] D. Sacha, H. Senaratne, B. C. Kwon, G. Ellis, and D. A. Keim. The role of uncertainty, awareness, and trust in visual analytics. 22(1):240–249. Publisher: IEEE.

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

[61] N. Sharma. Sensemaking handoff: When and how? 45(1):1–12. Publisher: Wiley Online Library.

[62] M. Skeels, B. Lee, G. Smith, and G. G. Robertson. Revealing uncertainty for information visualization. 9(1):70–81. Publisher: SAGE Publications Sage UK: London, England.

[63] P. J. Smith and R. R. Hoffman. *Cognitive Systems Engineering: The Future for a Changing World*. Crc Press.

[64] A. St Clair. Exploring the effectiveness of avalanche risk communication: a qualitative study of avalanche bulletin use among backcountry recreationists.

[65] G. Statham, P. Haegeli, E. Greene, K. Birkeland, C. Israelson, B. Tremper, C. Stethem, B. McMahon, B. White, and J. Kelly. A conceptual model of avalanche hazard. 90(2):663–691. Publisher: Springer.

[66] G. Statham, S. Holeczi, and B. Shandro. Consistency and accuracy of public avalanche forecasts in western canada. In *Proceedings ISSW*, pp. 1491–1496.

[67] D. A. Szafir, S. Haroz, M. Gleicher, and S. Franconeri. Four types of ensemble coding in data visualizations. 16(5):11–11. Publisher: The Association for Research in Vision and Ophthalmology.

[68] J. Thomson, E. Hetzler, A. MacEachren, M. Gahegan, and M. Pavel. A typology for visualizing uncertainty. In *Visualization and Data Analysis 2005*, vol. 5669, pp. 146–157. International Society for Optics and Photonics.

[69] J. Walny, S. Huron, C. Perin, T. Wun, R. Pusch, and S. Carpendale. Active reading of visualizations. 24(1):770–780. Publisher: IEEE.

[70] K. E. Weick. *Sensemaking in organizations*, vol. 3. Sage.

[71] D. Woods. Escape from data overload.

[72] K. Xu, S. Attfield, T. Jankun-Kelly, A. Wheat, P. H. Nguyen, and N. Selvaraj. Analytic provenance for sensemaking: A research agenda. 35(3):56–64. Publisher: IEEE.

[73] C. L. Yue, A. D. Castel, and R. A. Bjork. When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. 41(2):229–241. Publisher: Springer.

[74] T. Zuk and S. Carpendale. Visualization of uncertainty and reasoning. In *International symposium on smart graphics*, pp. 164–177. Springer.