

Open discussion <https://www.nat-hazards-earth-syst-sci-discuss.net/nhess-2018-183/#discussion>

Authors' response to [RC3 review](#) by M. Staples. The response is shown in blue.

1. Introduction

The challenges of communicating avalanche hazard to the public (section 1.2) were described very well. However, the first part of this section (lines 15-26) about impact-based warnings was confusing. If exposure and vulnerability are determined by individual users, how can AWS's issue impact-based warnings?

AWS issue regional forecasts, which describe the general probability and size of avalanches in a geographical area. These forecasts may describe the general impact for recreational activity, roads, buildings etc, but will not be able to determine the nature or scale of the impact on individuals or objects. For example, if an AWS forecasts a level 5 extreme danger in a mountain area where no one is recreating, the impact will be 0 as nobody are exposed to the danger. At the other side, the forecasts describes the impact to people that are at risk, if people choose to travel in avalanche terrain at extreme danger. So in a way the forecast is impact-based, but it cannot quantify the impact specifically as the AWS will not know how many people are exposing themselves to the risk.

We will rewrite the text to make it clearer, and possibly introduce an example.

U.S. successes (page 4, line 28) perhaps don't match other trends and could be worth mentioning. U.S system is unique in that it has very different styles and formats yet seems to be effective. The trend in the U.S. has been a declining fatality rate. The number of fatalities has been flat while use has surged, thus the rate has declined.

<https://avalanche.org/2016/06/27/2016627us-avalanche-fatality-trend-is-flat-for-the-past-22-seasons/>

Excellent point, we will write about this and refer to the article at

<https://avalanche.org/2016/06/27/2016627us-avalanche-fatality-trend-is-flat-for-the-past-22-seasons/>.

It would be very helpful to have an English version of Figure 2 (page 6). This would help some readers really understand the content of elements of avalanche warnings. Do any of the sections in the avalanche warning use stock language? Are they written from scratch each day? This is extremely important to know. Whether or not it contains original writing has a major effect on credibility.

We will translate the contents of the figure to English, we agree this would make it easier for the reader.

All text but that of the main message and the avalanche problems, is written manually by the NAWS forecasters. However, sometimes the forecasters may use parts of the text from the previous day, especially in the snow cover history. The text in the main message is produced in the following manner: The forecasters can get a list of text suggestions that are available depending on the chosen danger level and avalanche problem(s). He/she may then edit the text. He/she may also write the message from scratch or copy the text from the previous day. The text in the travelling advice in the avalanche problem is generated from a list of text suggestions. The selection of text is based on the combination of the chosen danger level and avalanche problem. The forecasters may edit the text afterwards. NAWS generate text suggestions in the forecast editing software for the main message and avalanche problems in order make the text in the warnings more easy for the users to read, ensure that the terms and wordings are as good and consistent as possible, make sure the time needed to produce the text is not too high and make translation of the text to English as good and easy as possible. NAWS have been having extensive discussions about to which degree should text be created from scratch by individual forecasters or rather

be predefined or suggested from a standard library of sentences and terms. NAWS is using a hybrid approach to this, and the creative from-scratch text is mostly found in the main message (Norwegian and English), avalanche danger assessment (Norwegian only) and snow pack and avalanche history (Norwegian only). However, creative text may suffer from poor language and significant individual difference that are difficult for the users to understand. Also, a lot of time may be spent writing text to convey a message that has already be written in much better way by someone else. However, the interest and motivation of users may drop if they get the feeling that too much text is auto-generated or copy-paste. NAWS is very aware of this effect and continuously make efforts to avoid this from happening.

We will explain how NAWS operate and elaborate the pros and cons of stock language versus creative text.

Many issues with this paper seem to stem from a translation issue. Use of some words like “danger” and “hazard” made this paper very confusing to follow. It was sometimes unclear when the word “danger” meant specific danger rating (i.e. Low, Moderate, etc.) or a general reference to hazard. Sometimes they called it the “danger rating” other times the “danger level”. While not a huge issue, this made the paper confusing.

We will revise the manuscript in order to improve this. We will make the use of “danger” and “hazard” more consistent or explain at the beginning that these terms are used interchangeably in the paper. The same applies to “forecast” and “warning” as well as “danger rating” and “danger level”.

Table 1 (page 7) needs to be rewritten. This table is referenced in Figure 5 (page 16); however, the element names in Table 1 and Figure 5 do not match. They are not presented in the same order either. Both of these issues make it very difficult to understand the results of this study. This was a serious problem for me.

Good point. We will fix this, by adding consistent names for each element in the table and the text.

Figure 2 shows elements of the avalanche warning. Each one is labeled with a number, and the caption has a description for each element. Figure 5 shows how users ranked elements in the avalanche warning. However, it was difficult to cross-reference these two figures because the wording in the caption for Figure 2 does not exactly match the wording in Figure 5. Additionally, the elements in Figure 5 are listed in a different order than they are in Figure 2. These issues made it challenging for me to fully understand the results.

Good point. We will fix this.

2. Methods and data collection

The way the male/female demographic was described on page 9, line 10 should be changed to mirror the way it is described in page 9, line 20. Line 10 perpetuates gender biases.

Good point. We will fix this.

Survey design: I do not know how to design a survey, but I know the way questions are phrased can have a big effect on responses. I assume that questions on this survey were written in a neutral way. It seems very difficult to truly assess comprehension. Given this difficulty, the authors did a great job trying to quantify comprehension with their system described on page 12, lines 1-5.

Thanks!

Why did you test comprehension of wet slabs? Do they kill a lot of people in Norway? Persistent slab avalanches kill many people in the U.S., and they are difficult for users to understand. Risk management and travel advice messages for persistent slab avalanches are difficult to

communicate. Additionally, there can be significant message fatigue with this avalanche type. Personally, I would have tested comprehension regarding this avalanche type.

We wanted to test different danger levels and different avalanche problems. However, we had to reduce the scenarios to two. We agree that testing a persistent dry slab problem would be interesting. However, the wet slab problem is a relevant problem in Norway when considering high danger levels and the danger of natural avalanche.

We agree to the suggestion to make another study, in order to test more avalanche problems. It would also be useful to test more specific comprehension alternatives, and how these translate into real-life management of the avalanche threats in the field.

Wind slabs were a good problem to test because they are so common.

We agree, and this is a very common problem in Norway.

The four alternative ways to present the forecast (is it a forecast or is it an avalanche warning?) on page 11, lines 10-12 do not match items listed in Table 2, section D. For example, item 1 is listed in the table as "Avalanche danger with explanation (general advice associated with the danger level)" and it is listed on page 11 as "only the avalanche danger level and very limited travel advice". This is not a major issue, but it makes it hard for me to follow the paper. Is "general advice" the same thing as "very limited travel advice"?

Good point. We will fix this.

The communication effectiveness score, page 13 line 1, seemed like a great way to assess the responses from participants. Is it perfect? Who knows? Using "expert" answers as a way to evaluate participant answers seems like a great process to me.

Thanks! It is never perfect, but we thought it would be a great way to measure if the users grasped the same message as the people working for NAWS tried to convey in the warning.

3. Results

Another inconsistency involved the level of avalanche knowledge. The categories mentioned on page 13, lines 18-21 do not match the categories of "competence" listed in Table 4 (page 14).

Good point. We will fix this.

In table 4 the level of experience is categorized by "ski tours per year". How did you measure the experience of other users like snowmobilers, snow shoers, etc?

Good point. We will fix this by replacing "ski tours" with "tours in avalanche terrain". This is a translation mistake, as the question asked in the survey was independent of how the trip was carried out: "Travel in avalanche terrain - during the winter do you do 0, <5, 5-15 or more than 15 trips in avalanche terrain?".

Again, the warning elements in figure 5 do not match the elements of the warning listed in table 1 and figure 2. While the authors may be referring to the same elements, using different wording to describe the elements made it very difficult for me to understand.

We will fix this.

The authors did a great job summarizing the qualitative results. In the few surveys I have conducted with users in the U.S., comments often contain the most valuable info. Sometimes a single comment from a single person can be the most valuable part of the survey.

We agree.

Page 21, lines 9. The authors comment that “user’s competence had no effect on the ranking” in line 9. Did experience have an effect? As I understand it, “competence” and “experience” are two different things.

Yes, experience is also non-significant. The F value for competence was incorrect; will be corrected to 1.966:

A user’s experience had no effect on the ranking of the alternatives, $F(1, 172) = .469$, $p = .494$, $\eta^2 = .002$.

A user’s competence had no effect on the ranking of the alternatives, $F(1, 172) = 1.966$, $p = .163$, $\eta^2 = .010$.

Similarly, in line 20, the authors comment that “experience did not influence the ranking.” What about competence? Later in line 25 they say that “competence had no effect on comprehension”. Which is it? Is it both? This is confusing and not clear.

Also user’s experience did not influence the ranking, $F < 1$.

We will change to: Also user’s experience and competence did not influence the ranking, both F ’s < 1 .

4. Discussion

Page 22, lines 20-22 – That sentence states the purpose of an avalanche warning really well.

Thanks!

The discussion of symbols vs text, page 23 line 25, is interesting because the tech industry has struggled with and gone back and forth on. Symbols can be confusing. Text is not, but there are issues in translating between languages.

We agree.

Page 24, line 15, The words danger and hazard seemed to have been used interchangeably. While they may mean the same thing, it would be better to pick one for this paper. It could be easier to use the word “danger” when referring to the danger level. This added an extra layer of confusion for me.

We will revise the manuscript in order to improve this, see comment earlier.

Page 24 lines 22-24 suggest that danger level is well understood by users. Section 3.2.3 suggests that users have difficulty understanding the danger level. Another inconsistency is that the authors would sometimes say “danger level” and other times say “danger rating”. It would help to use just one.

The first point we do not consider a major discrepancy, our interpretation is that it is rather difficult to understand/respond to a level 2 warning while a level 4 is more easy to understand/respond to. We will revise the manuscript with respect to the use of “level” and “rating”, see comment earlier.

All of page 25 is a great discussion. A huge issue is you addressed is when 2-3 avalanche problems are present. There are almost always 2 problems present. Great topic for further discussion in another study.

Thanks! We agree that this would be a nice topic for further studies.

5. Conclusions and recommendations

This study delivered some concrete findings. It seems to have been well designed, but it was challenging to understand because of the writing. The authors did a good job with their

conclusions by not presenting conclusions with too much specificity. It would be easy to read too much into the results, but they did a good job of keeping their conclusions more general.

Thanks! Your comments will be of great value for revising the manuscript.