

Reply to reviewer 2: Anonymous

Title: Key learning moments as predictors for understanding snowpack dynamics during a season-long avalanche course?

Author(s): Tim Dassler, Richard Fjellaksel, and Gerit Pfuhl

MS No.: egosphere-2024-1533

MS type: Research article

Special issue: Latest developments in snow science and avalanche risk management research – merging theory and practice

Dear Reviewer 2, dear editor,

We are immensely grateful for your time and constructive feedback, which we have used to improve the structure and clarity of the manuscript.

Please find below our reply to your comments.

Overview of revisions:

Here we provide a brief overview of the major revisions made to the manuscript based on your comments.

1. We changed the title and included the word 'case study' to address more clearly the characteristics of our study.
2. We reduced the focus on key learning moments (e.g. removed in title, provided simple definition incl. examples).
3. We substantially restructured and rewrote the introduction, methods, findings and discussion section to improve structure and clarity.
4. Suggested literature was included in the discussion section.
5. To make the language more inclusive of the general backcountry population we changed 'safe to ski' assessment to 'safe to descend'. Similar changes were made to the language throughout the paper.
6. We have prepared a revised draft, which you will find in the folder 'Revision_Draft_SafeToDescend' at the following location: https://osf.io/26x5z/?view_only=02d73beecc9a404bb79f3588b80f5469
7. Our line and section numbers refer to the revised version.

Point-by-point response:

Below we address each comment raised individually.

Reviewer comments are marked in bold font, our replies in regular font.

Overall comment:

Avalanche education is an understudied field, and the article offers a contribution to our understanding of the efficacy of avalanche courses. The third research hypothesis is clever; approaching education as an opportunity to establish one's own knowledge gaps is brilliant. There is much value in this train of thought, including the authors' recommendations for providing repeated and novel opportunities to life-long learning for mountain travelers (L611).

Despite the potential benefits of this paper, I have some fundamental concerns about the employed methods and the interpretation of the quantitative results. In my opinion, addressing these issues requires a substantial repositioning and rewriting of the manuscript before it can be considered for publishing.

Here are the topical comments on the manuscript focusing mainly on high level points about the methodology and results; I am not providing line-by-line comments for this version.

Main concern is the limited data set:

The dataset is very small (8-10 participants) for the application of linear mixed models or Chi-square tests (L279). The number of key learning moments in the dataset is neither included in the article nor the supplemental materials. The authors applied Cramer's V, that is less sensitive to sample size, and pivoted to using effect size when discussing the results. Yet it seems likely that the data set is not viable for Chi-square testing and that extremely small sample can influence the strength of associations used in the results. The results may not be generalizable.

We thank you for challenging our analysis. Statistical power for repeated measurements is higher for smaller samples than for a similar sized cross-sectional study. As replied to reviewer 1 in four of the outcome variables the effect sizes were large enough that a power analysis yielded over 80% power.

We are aware of the limitations of Cramer's V, however a major advantage of Cramer's V is that it represents the association between two categorical variables as a number between 0 and 1, similar to correlation coefficients and likely easier to understand for most readers (provides familiarity as already Cohen himself noted as a drawback for using Cohen's omega). We cannot use the Phi coefficient as we have larger than 2x2 tables.

If the article's objective is to share the quantitative results (L19-21; section 2.3), this goal is out of reach due to the limited sample size. The present focus on a quantitative analysis seriously distracts and devalues from more valuable qualitative insights that the study could provide. Hence, reframing the inquiry as the exploration of qualitative insights from the survey data would be more appropriate objective for the manuscript considering the available data set.

We clarified that this is a case study, e.g. changed the title, and that the focus of this paper is the relation between 'subjective' experience of learning (through key learning moments) and 'objectively' measured learning outcomes (snowpack assessment skills and applied understanding). Since we used an exam-like assessment (AviLog) to measure understanding of snowpack stability, a quantitative approach is justified. We do not claim that our findings are generalizable to the broader winter backcountry population. We discuss that our group of participants also was a highly motivated group.

We further disagree that we cannot report quantitative results given our sample size. Our study is designed as repeated measurement and has more power than a similar sized or larger sized study with two independent groups. A sample size of $n=10$ has over 80% power to detect large effects (Cohen's $D \sim 1$ or larger). See also <https://journalofcognition.org/articles/10.5334/joc.72> and <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6701714/>

It is correct that we did not perform an a priori power analysis, as we had no published effect sizes on which to base the power analysis. However, we did expect large effects from baseline to module 4. We had no prediction for the follow-up one year later. Since we have not pre-registered our hypothesis, we did not include a power analysis.

As mentioned in the manuscript, we do not analyze the qualitative data sets (field observations and focus groups) in this study as this beyond the scope of this paper and will be part of a forthcoming publication.

Connecting the dots was difficult:

The lengthy introduction discusses foundational experiential education literature combined with more current references on effective and transformational learning theories and a short list of relevant articles on avalanche education research.

Notably missing is McNeil et al. (2023) that addresses topics of L53-55. Editing the introduction to focus on major theories related to key learning moments, knowledge acquisition, and the role of snowpack information in avalanche decision-making, would prioritize the content that is most relevant for the results and discussion.

It was difficult to follow the connection from hypotheses to conclusions. This may be due to the unclear explanation of the results or the verbose and ambiguous language. If the connections between research questions, theory, data, analysis, and results would tie together more coherently, it would be much easier for readers to understand the conclusions and implications of the study and how well they are grounded in theory and the collected evidence.

We restructured and shortened the introduction for better readability focusing on theories relevant to key learning moments, knowledge acquisition, and the role of snowpack information in avalanche decision-making.

We also clarified the connection between research questions, hypotheses and results and addressed the relation between the experience of learning (key learning moments) and measurement of snowpack assessment skills (AviLog, 'safe to descend' question).

In the introduction we make a point that there are few peer-reviewed studies on avalanche education. But we included the suggested reference (together with other similar studies) in the discussion of our findings.

Key learning moments

Here are three examples that create convolution related to the title concept, key learning moments (KLM) (Section 1.3), the metric the authors emphasize as a major factor in the efficacy of education:

- 1. There is no data of KLMs included in Section 2.1 or elsewhere in the paper.**
- 2. In the results section, statements on the variable associations with KLMs are hard to understand (L348-353).**
- 3. X axis on Figure 3 (L370) is labeled "average KLMs", but it is unclear if this number is average KLMs per trip or an overall mean of KLMs per participant during the study period; a clear caption would clarify the figure interpretation.**

We agree and thank you for your comments to reduce convolution. We edited the introduction to provide a more coherent definition of key learning moments, including examples, and their importance for experiential learning.

Method section 2.2.1 on key learning moments was revised for clarity.

In the results section findings related were clarified regarding Hypothesis 2 and that the amount of key learning moments increases the request for support, but had no significant association with subjective or objective competencies of assessing snowpack stability (L418-420).

'Why to ski'

Another mind bender was the 'why to ski' construct the authors use to measure the participants reasoning to decide that it was safe to engage in the activity. The 'why to ski' score is introduced in section 2.2 (L267) without a transparent connection to research question about students' ability to apply their knowledge on their own (L154) or their confidence to justify their field decisions (L164).

Important point! Thank you. We clarified the connection between the 'safe to descend' (previously 'why to ski') question and measuring applied understanding of snowpack assessment in the introduction.

We also clarified how we scored participant answers to the 'safe to descend' question based on factors from the systematic snow cover analysis (presence/absence of weak layers, initiation/propagation of cracks, depth of weak layer, thickness and hardness of layer above, etc.)

On a different note, the focus on the term 'ski' perpetuates the imbalance in the representation of various mountain activities in avalanche safety research. Even if the dataset consists solely of skiers, the researchers could use more inclusive language when presenting their findings.

This is an important point and something that slipped our attention. Thanks a lot for catching this. We changed 'why to ski' into 'safe to descend' as well as 'ski' and 'skiers' to other terms such as 'backcountry recreationists' 'backcountry travelers' etc. throughout the text.

Snowpack analysis lessons

An additional content gap is the missing information on what was taught about snowpack analysis. The authors refer to standardized curricula from Norwegian Mountain Forum (L209) but also that the participants had their say about the content (L157). Including the actual snowpack topics covered, practiced, and reviewed in each module would be informative as an appendix to explain the educational delivery of 23 snowpack factors measured (L164) and analyzed (L260).

This is a valid point. We added a new section (2.1.2) where we describe how snowpack analysis was taught during the course. Here we also present the learning goals for snowpack analysis which were the template for scoring the why is it 'safe to descend' answers. We also added an info box (Figure 1) that explains the Systematic Snow-Cover Analysis taught during the course.

Contributing factors outside the course participation:

Factors outside the participation in an avalanche education experiment can contribute to the scores in your data; for example, learning moments that happen on personal tour days can build competency. It is possible that the participants engaged in active experimentation outside the course modules (L403). Were these data (L287-293) collected from the participants repeatedly or only once for a baseline?

We agree that learning moments and factors outside the course can influence learning and build competency. Participants were also encouraged to explore and experiment outside the course modules as this is a cornerstone of experiential learning. This was done e.g. through 'homework' assignments. See newly added section 2.1.2. However, our main findings indicate that while course participants 'objectively' improve in their snowpack assessment skills during the course (which may be attributed to the course content or other factors), this improvement is not lasting for all a year later. We actively address external factors which we did not measure that might provide alternative explanations of these findings, including the lack of authentic learning experiences during and lack of active experimentation after the course in the discussion section.

And more generally, did the authors consider how to contextualize the results into the participants' individual lived experiences – not only as participants of four education modules? There is much potential for deep descriptions on the individuals over the repeated interactions with ethnographic approach (Dassler et al., 2023), but that rich viewpoint is missed in this paper. By reframing the paper, the authors could provide a more detailed introduction of their participants, highlighting them as key players in the case study.

We wholeheartedly agree that the qualitative dimension of learning experiences deserves proper treatment. Since we are here concerned with the analysis of learning outcomes related to snowpack assessment skills regarding their relation to experienced learning moments, we would not be able to do justice to the rich experiences and data collected. This will be the focus of a forthcoming publication that focuses on the qualitative dimensions of the data.

We substantially restructured introduction, methods, findings and discussion to make focus and aim of the paper clearer.

Explanation of how all the parts of the project fit together:

As a part of a larger research project, the article made references to other papers related to the education experiment (i.e. L135 and 165), which was slightly distracting for this specific piece. I needed to locate and read the other articles mentioned in the text to get the full understanding.

We highly appreciate the extra work invested. By restructuring our introduction, the distractions should now have been vanished.

It would have been helpful to have a concise reference to a holistic framework of how the different pieces fit together, perhaps as a flow chart or a table rather than the introductory paragraph in Section 2.2. Sharing the necessary information from the adjacent articles clearly and concisely would improve the thread of the manuscript.

Agreed. We reduced the need to read external articles by providing a rewritten section to the question how we measure learning (subjectively and objectively) in the introduction. This section includes a simpler and more coherent definition of key learning moments and why they are important for learning.

We restructured and rewrote the introduction to improve coherence and clarity. Among other things, we provided a clearer and simpler definition of key learning moments and how we used them in the study.

We included a new figure and a table. Figure 2 shows a timeline of the modules, participants and when the AviLog was filled out (objective measure of snowpack assessment skills). Table 1 is an overview of data collected in each module.

Conclusion:

Dear Reviewer 2, we highly appreciate the expertise, time and effort you put into providing constructive feedback. We believe that by addressing your concerns the manuscript has been improved considerably, both in structure, clarity and readability.

We hope that the revisions meet your expectations and concerns.

Kind regards,

Tim Dassler, Richard Fjellaksel and Gerit Pfuhl