

SKIING IN AVALANCHE TERRAIN WITH OR WITHOUT A LEADER: HOW DOES THE SKIERS' EVALUATIVE CRITERIA FOR GROUP PERFORMANCE DIFFER?

Rong Guang¹, Lauri Ahonen¹, Gerit Pfuhl^{2,3}, Audun Hetland³, Benjamin Ultan Cowley^{1,4,*}

¹Faculty of Educational Sciences, University of Helsinki, Helsinki, Finland

²Department of Psychology, NTNU, Trondheim, Norway

³Center for Avalanche Research and Education, UiT the Arctic University of Norway, Tromsø, Norway

⁴Cognitive Science, Faculty of Arts, University of Helsinki, Finland

ABSTRACT:

A competent leader can help to improve group performance, for example by helping group members to choose actions that reduce risk; however an incompetent leader can hurt performance and increase risk when group members follow uncritically. Recreational ski groups rarely have a formal leadership hierarchy. Here, we conducted a study to investigate skiers' self-evaluation of touring in groups with and without a leader. We developed a survey measuring skier' group dynamics, interpersonal trust and confidence in decisions. We surveyed 215 skiers (two cohorts: with-leader, n=101; without-leader, n=114) about their last ski trip. We tested between-cohort differences on group performance. Then we used factor analysis to reduce data dimensions of each cohort separately, to find components that reflect what the skiers found important in group performance. A group of experts discussed and agreed upon the nature of components. The tests did not reveal any differences in trust and confidence between cohorts, even after controlling for potential confounders. However, we found skiers' evaluative criteria substantially differed between cohorts. Specifically, skiers with a leader emphasized the role of group openness, followed by members' relevant competence in avalanche assessment and then a sense of involvement; those without leaders highlighted preparedness, followed by harmony. Both match logically with leader presence/absence, and criteria for without-leader skiers are arguably characterised by less sensible group risk mitigation strategies.

Keywords: Backcountry skiing, leadership, group dynamics, group performance, evaluative criteria

1. INTRODUCTION

Backcountry (BC) skiing very typically happens in groups (e.g. 91% of backcountry skiers surveyed by Procter et al. (2014), and 90 of 105 BC skiing 'groups' surveyed by ourselves (Ahonen et al., 2024), had ≥ 2 members). Indeed, a social component has been identified as an important aspect of human factors of decision making in avalanche risk management (Ebert and Morreau, 2023). Despite the established advantages to being in a group with a leader, such as enhancing group cohesion (Light Shields et al., 1997), there is, however, little research that investigates why BC skiing groups often travel in absence of a formal or informal leader. This article aims to shed light on the topic.

BC skiing is a popular yet dangerous outdoor recreation (Birkeland et al., 2017). BC skiers must navigate through snow-covered, steep mountainous areas that present multiple risks, including fatigue, getting lost, accidental falls, collisions with terrain obstacles, and avalanches (Ferguson and

LaChapelle, 2003; Page et al., 1999). Over 60% of BC skiers tour in groups of 2–5 people (Procter et al., 2014), with groups of 3–4 believed to be the safest (Tremper, 2013). Although there is observational evidence showing solo skiing is safe (Zweifel et al., 2016), expertise can be a confounder (solo skier might have stronger skills). Safety recommendations advocate for group travel (Tremper, 2013; Ferguson and LaChapelle, 2003). Indeed, group members can share situational awareness (Salmon et al., 2010), facilitate decision-making (Tremper and Diegel, 2014), monitor each other for signs of fatigue (Grosse et al., 2007), and provide assistance in case of accidents (Wallner et al., 2019).

However, a group could also have side effects on decision quality. For example, decisions in a group with maladaptive dynamics, like groupthink – where a group under stress may develop a false sense of cohesion – can amplify ignorance or incompetence of group members (Ebert and Morreau, 2023). In Lewin (1943)'s group dynamics theory, groups are influenced by both opportunities and threats, which can strongly affect their collective behavior.

Specifically in BC skiing, McCammon (2002) identified the social proof heuristic ¹ and the expert

*Corresponding author address:
B. Cowley, University of Helsinki
PL9, 00014, Helsinki, Finland;
tel: +358 50 3199 335;
email: ben.cowley@helsinki.fi

¹tendency to believe that a behavior is correct to the extent that other people are engaged in it

halo² as common traps. The findings are supported by empirical studies (Couret et al., 2020; Hendriks and Johnson, 2016) and avalanche accidents where poor group dynamics can be the contributing factor (Ebert and Morreau, 2023).

Based on observations and interviews of skiers, Zweifel (2015) suggested six topics as key for good group dynamics: skills, organization, communication, identification with common goals, anomalous relationships, and leadership. Among these factors, leadership has long been implicated in shaping group dynamics (Baumeister et al., 1988). A perceived expert's (leader's) judgment is nonetheless the ceiling of collective decision in a negative case scenario when a group of skiers blindly and thoughtlessly follow the leader's decision (Ebert and Morreau, 2023). Data collected by the authors (at a single ski site in 2023) found that 43% (96/221) of BC skiers had their trip without a formal or informal leader (Ahonen et al., 2024), illustrating that the dichotomy between groups with or without identified leaders is meaningful.

Assuming skiers seriously consider the high stakes of each trip, they should have different criteria for perceiving group dynamics to compensate for leader absence. Several theoretical frameworks have potential to explain such difference, such as social identity theory (Kerr and Jermier, 1978) and shared leadership (Zhu et al., 2018). However, the mechanism underlying the phenomenon has yet to be clarified due to a lack of empirical evidence. Clarifying this mechanism can benefit the community to help: understand group behavior, improve teaching programs, develop safety recommendations, and enhance group performance. Here, we investigate whether and how such adaptation occurs, answering two research questions:

1. Do skiers without a leader differ in their self-perceived group performance compared to those with a leader? (RQ1)
2. Do they differ in the criteria for self-evaluating group dynamics, and if so, how? (RQ2)

2. METHOD

2.1 Participants and procedure

We used a sample recruited from the CARE panel and an avalanche course. The CARE panel (Mannberg et al.) is run by the Centre for Avalanche Research and Education (CARE) at the Arctic University of Norway, Tromsø (UiT), and provides a non-representative cohort of over 2,000 adult backcountry adventurers. All panelists received an invitation

to join the study and participated voluntarily. No factors influencing their participation are expected to be associated with group performance.

Upon entering our online survey, respondents read an introduction to the study, estimated time of completing the survey (<10 mins), and an informed consent. Respondents were rewarded with entry to a prize draw. Through the survey, we collected respondents' demographics and, according to the most recent BC trip with a group that they were asked to recall, we further collected the presence of formal/informal leader, trip familiarity, group dynamics, trust with group members (referred to as trust herein) and confidence in decisions made (referred to as confidence herein). Withdrawal from the study was allowed at any point before data analysis. Upon completing the final questionnaire, participants were thanked for their time and contribution. To reach the required sample size, we repeated the process with participants from an avalanche course in 2023. Data collection was approved by the Norwegian Agency for Shared Services in Education and Research (SIKT, formerly NSD 733888).

2.2 Measures

We aimed to measure group performance (meaning the quality of group interactions, not their performance in any objective skills) and individual skills. We adopted the well-established group-inputs, -processes and -outcomes framework (IPO) (Kendall and Salas, 2004) to evaluate group performance. Group inputs are those individual and group characteristics that serve to inhibit or facilitate group processes such as coordination, communication, and decision-making (Kendall and Salas, 2004). Group processes refer to how the group attains its goals (Rosen and Dietz, 2017), or specific types of behaviors leading to desired group outcomes (Kendall and Salas, 2004). In contrast, group outcomes are tangible and intangible results (e.g. completion time, or emotional payoff) of group members' interactions and joint processes (Kendall and Salas, 2004).

2.2.1 Group Processes

Short-term group dynamics are an approximate indicator of group processes. To evaluate group dynamics for BC skiers, we adapted the SOCIAL tool, which was developed specifically for the BC skiing context by Zweifel (2015). It is a 20-item ski-group dynamics checklist, which include six normative topics: Skill, Organization, Communication, Identification, Anomaly and Leadership, each having 3 to 4 items. It presents a comprehensive list of questions suggesting good practice in skiing group interactions (e.g., "Would everyone voice his concerns at any time?"). SOCIAL was developed initially as a group check tool in order to improve the

²tendency to defer all judgment to leader of the group

control of group dynamics within winter BC groups. Via an expert panel procedure, we adapted it into a Norwegian instrument (referred to as SOCIAL scale herein), with responses for each item captured on a 5-point scale anchored at 1 (strongly disagree) and 5 (strongly agree). Following an established procedure (Roberts and Thatcher, 2009), the instrument was further validated as a formative scale (see <https://osf.io/wqx2j>). One item from the initial SOCIAL tool was removed as the result of the validation (“Are any love stories going on in the group?”), resulting in 19 items.

The formative nature of the validation allows to keep the initial six components, which are crucial aspects of a group process (Hanafiah, 2020). Since not all groups have a leader, the survey was designed to refrain from showing three leadership items for respondents who reported they had no leader on that trip.

Scores of the items were averaged such that higher scores reflected higher group dynamics. Scores for items under each of the six components were averaged to reflect quality of each component.

2.2.2 Group Outcomes

To obtain subjective group performance of skiers, we further included measures of group outcomes. Since avalanche is a low-probability high-stake event (Page et al., 1999), a trip without an accident does not necessarily imply acceptable group performance, such that tangible results might not be the optimal measure. Many intangible constructs have been linked conceptually with performance and have been assessed at the individual level, such as trust (De Jong et al., 2016), group cohesion (Casey-Campbell and Martens, 2009), and collective efficacy (Kendall and Salas, 2004). We chose to measure trust for our study, considering:

- (a) Trust is an intuitively understandable concept, and can be asked with one item (Castro et al., 2023). It helps keeping survey length short and simple, which helps response quality (Galesic and Bosnjak, 2009);
- (b) Trust may work better in a self-report survey, in contrast to cohesion and collective-efficacy which might require observers, e.g. to provide subjective ratings (Kendall and Salas, 2004);
- (c) Trust may be more important than cohesion for groups whose goal can be disjunctive (Steiner, 1972) such that the most proficient members in avalanche assessment can (potentially) have full weight in group performance;
- (d) Common scales for group cohesion (Wongpakaran et al., 2013) and collective efficacy (Carroll et al., 2005) have overlapping items with the SOCIAL scale.

As such, trust was scored on a 5-point scale from 1 (minimal) to 5 (complete trust).

Additionally, evaluations of group performance vary by group type and goals. Hence, the indicators must be defined contextually (Kendall and Salas, 2004). Skiers need to make critical decisions such as go/no-go for a slope, and uncertainty (confidence) in such decisions can be linked to performance. Our earlier field study used the ideas of “critical decision point” and “confidence in critical decision point” in its surveys. Face validation showed skiers understand the concept properly (Ahonen et al., 2024). Borrowing the concept, we included one more group-outcomes indicator – confidence in the decision made.

Confidence was scored on a 7-point scale from 1 (minimal) to 7 (complete confidence). The scales were set with varying numbers of points to avoid common response bias.

2.2.3 Moderator: trip familiarity and contribution

Perception of some team constructs can be influenced by the individual rater (Schmidt et al., 2023). Members role may have different perception of the necessity for vigilance to avalanche cues, when they have done the trip several times (McCammon, 2003), or when they have a special role in making group decisions (Kolb, 1995). As such, we measured the respondents’ trip familiarity and contribution to decisions, as moderator of group performance.

Trip familiarity is collected by an item asking “How many times have you done this trip before?”. Options include “Never, this is the first time”, “Once”, “2–3 times”, “4–10 times” and “More than 10 times”. Contributions of the respondent in the group were obtained by asking “What was your role in decision-making on this trip?”. Options include: a. “I made all the decisions on my own without consulting with the group”; b. “I made the decisions but I consulted with the group”; c. “I was part of the discussions that led to the decisions”; d. “I had an opinion but did not voice it”, and e. “I had no opinion and just followed the group”. They were further re-coded as ‘High contributor’ (a+b), ‘Medium contributor’ (c), and ‘Low contributor’ (d+e).

2.3 Data analysis

We took measures to improve data quality. For details, see <https://osf.io/762cv>. In the analysis, respondents were grouped into with-leader and without-leader cohorts according to the self-reported presence of a leader.

To answer RQ1, we used regression analyses to investigate inter-cohort (with- vs without-leader)

difference in processes (including overall group dynamics and each of its six indicators) and outcomes (trust and confidence). By further investigating the difference between each of the six indicators, we can identify those playing essential roles. Recognizing no single measure can comprehensively capture group performance (Kendall and Salas, 2004), we complemented by further investigating psychological payoffs from the trip (outcomes): trust and confidence. Considering the role of individual's skill in influencing their self-reported group performance, all these analyses were done by controlling for trip familiarity and contribution to decisions.

For RQ2, we did a factor analysis on group dynamics items to identify latent factors for each cohort, respectively. Here factor analysis was *not* used for establishing factor validity. Instead, we used it to look into the respondents' 'mechanism of preferences' underlying findings from the above regression analyses. This is possible since any pattern in self-reported group performance should be generated through implicit evaluative criteria harbored by the skiers, and it serves to identify them.

2.3.1 Analytical Procedure

Baseline characteristics, including gender, age, trip familiarity, and contribution to decisions, were compared between cohorts to ensure comparability.

Averaged overall group dynamics, each of its six factors, trust, and confidence were compared between cohorts. Generalized linear regression was used to address the ordinal nature of these variables, with the link function determined based on variable distributions. This method was chosen for its ability to model ordinal outcomes accurately. Sample characteristic that differed between cohorts were controlled in the regression analysis.

For baseline comparisons, an alpha level of 0.05 was used to assess inter-cohort equivalence. For primary analyses, Bonferroni correction was applied, adjusting the *p*-value threshold to 0.004 (0.05/12) to account for multiple comparisons.

Factor analysis was conducted on group dynamics items for with- and without-leader cohorts, separately. Factorability was checked by KMO and Bartlett's test (KMO > 0.5 and Bartlett's test *p* < 0.05 were considered acceptable). Principal Axis Factoring (no multivariate normality is assumed) was adopted as factoring method. The optimal number of components were selected based on the Scree plot and Very Simple Structure. Orthogonal rotation was used. The process is demonstrated in a flow chart, see <https://osf.io/6wa3p>.

Results of the factor solutions were compared against each other. One solution was kept as final solution when *a*. it does not include single-item factor and *b*. it explains the highest amount of variance.

Table 1: **Sample characteristics by cohort**

Characteristics	With-leader cohort (n = 100)	No-leader cohort (n = 115)	<i>P</i>
Mean age <i>y</i> , (SD, <i>n</i>)	36.6 (11.1, 85)	35.5 (10.3, 97)	0.27
Male, <i>n</i> (%)	72 (93.5)	84 (90.3)	1
Trip familiarity, <i>n</i> (%)			0.54
Never	49 (49.0)	50 (43.5)	
1-3 times	28(28.0)	31(27.0)	
>4 times	23 (23.0)	34 (29.6)	
Role in group, <i>n</i> (%)			0.03
< high contributor	70 (70.0)	96 (83.5)	
High contributor	30 (30.0)	19 (16.5)	

The selected factor solutions were deemed as skiers' evaluative criteria, with each factor indicating one criterion. The preference for each criterion was ordered by the amount of variance explained by the factor (criterion). We presented the factor structure to a panel of experts. They discussed, agreed upon, and assigned names to the nature of components. We qualitatively compared the results by comparing the existing evidence.

2.3.2 Sample size

We aimed for at least 95 participants for the with-leader cohort and 80 for the without-leader cohort. The calculation was based on Cohen (1988) and Bryant and Yarnold (1995). For details, see <https://osf.io/8pgtd>

3. RESULTS

3.1 Sample characteristics

Among 277 skiers who were surveyed, 215 (78%) valid responses were achieved. Appendix 'data quality control flowchart' presents a flowchart.

The valid sample consisted of 101 respondents whose group had a leader; 114 respondents having no leader. 159 (94%, available *n* = 170) were male. Mean age was 36 ± 10 (available *n* = 182). Ninety-nine skiers (46%) had never travelled the same route before the trip (available *n* = 211). Ten skiers (5%) were low decision contributors (available *n* = 215). Considering they were extremely under-represented in the data, we re-coded low and medium decision contributor groups as '< high contributor' in the following analysis.

Table 1 presents sample characteristics. The with-leader cohort had more high-contributors. No differences were revealed for other characteristics.

3.2 Generalized linear regression

Generalized linear regression (with gamma link function for positive skew) was used to study the role of leader presence in group performance. For easier interpretation, we exponentiated the estimates. An exponentiated coefficient >1 indicates an increase

Table 2: Generalized Linear Regression of Group Performance Indicators

Predictors	GD	GD components					Trust	Confidence
		Skill	Organization	Communication	Identification	Anomaly		
With-leader	0.94 * (<0.001)	0.86 * (<0.001)	0.97 (0.226)	0.98 (0.447)	0.97 (0.244)	0.99 (0.590)	0.99 (0.568)	0.99 (0.750)
High contribution	0.97 (0.147)	0.82 * (<0.001)	1.02 (0.561)	0.97 (0.243)	1.01 (0.835)	1.01 (0.581)	0.99 (0.634)	0.98 (0.554)

Note: GD: Group Dynamics. Reference for predictor "With-leader" is "Without-leader"; Reference for "High contribution" is "lower contribution than high"

in the expected outcome, while an exponentiated coefficient <1 indicates a decrease.

Trust ($\exp(\beta) = 0.99, p = 0.568$) and confidence ($\exp(\beta) = 0.99, p = 0.750$) were not affected by the presence of a leader. The presence of a leader was significantly associated with overall group dynamics ($\exp(\beta) = 0.94, p < 0.001$), indicating that with-leader skiers tended to report slightly lower overall group dynamics compared to those without a leader. Specifically, the skill component was significantly lower in groups with a leader ($\exp(\beta) = 0.86, p < 0.001$). No significant differences were observed for other group dynamics components (see table 2).

3.3 Factor analysis

3.3.1 With-leader cohort

For the with-leader cohort, after removing one item which correlated 0.3 with any other item, KMO measure was 0.713 (Good acceptance). Bartlett's test was significant ($\chi^2(153) = 483.67, p < 0.001$), suggesting redundancy among variables that could be summarized with fewer factors. Hence, factor analysis was deemed suitable with the 18 items.

Factor analysis on these items indicated a three-factor solution as easiest to interpret, and explained most variance at 62.4%. The two-, four-, and five-factor solutions explained 44.8%, 57.5%, and 60.8% of the variance respectively, and had various issues such as negative loadings. The three factors in the chosen solution explained 32.2%, 20.5%, and 8.7% of variance of the items (see table 3).

3.3.2 Without-leader cohort

For the without-leader cohort, after removing two anomaly items which correlated 0.3 with any other items, KMO measure was found to be 0.76 (Good acceptance). Bartlett's test was significant ($\chi^2(105) = 448.005, p < 0.001$), suggesting redundancy among variables that could be summarized with fewer factors. Factor analysis was then deemed suitable with 14 of the 16 items.

Factor analysis on the items indicated a two-factor solution as easiest to interpret and explaining 38.2% of the variance. Although the three-, four-, and five-factor solutions explained a higher amount of the variance, they were not chosen due to issues of single-item factors and/or lower interpretability. The two factors explained 25.7% and 12.5% variances of the items, respectively. See table 3.

3.3.3 Expert panel

The resulting best factor solutions were presented to a panel of avalanche researchers (n = 2), forecasters (n = 3), and behavioural scientists (n = 3) in two rounds, each having four experts. They agreed on the nature/name of the three factors for the with-leader cohort (in decreasing order of variance explained): *Openness, relevant competence, and sense of involvement*; and on the nature/name of the two factors for the without-leader cohort: *Preparedness and Harmony*. See table 3.

4. DISCUSSION

4.1 Primary finding

We found the absence of a leader does not significantly lower skiers' perceived group performance: they perceive same group-outcomes and better group-processes than skiers with leader. However, this is the result of their different evaluative criteria in measuring group performance. Skiers with leader focus on healthy interactions and avalanche assessment skills; whereas skiers without leader focus on preparedness and harmony.

4.1.1 Answers to research questions

To answer RQ1, we regressed skiers' self-reported group-outcomes and group-processes on if they have a leader. The results did not reveal a difference in trust and confidence between with- and without-leader skiers. It is noteworthy that they did differ in the group-processes indicator-group dynamics and its skill components, suggesting a different evaluative criterion might be working that led to the comparable group-outcomes.

To answer RQ2, we did factor analysis on group dynamic items for with- and without-leader skiers, respectively. The result suggests skiers with a leader emphasize, by decreasing order of importance: *openness, relevant competence, and sense of involvement*; whereas those without a leader emphasize *preparedness* followed by *harmony*.

4.1.2 Different evaluative criteria

A few interpretations to the name/nature of evaluative criteria are possible. Our most agreed explanation is: With-leader skiers know that it is the leaders decisions that will affect them. Therefore

it makes sense that they are interested in openness (factor 1) such that all skiers can follow along; then avalanche assessment skills (relevant competence, factor 2) to contribute to decisions; and finally, that they are included in decision-making (sense of involvement, factor 3). In contrast, without-leader skiers cannot rely on a leader for coordination and need to do all the work themselves, therefore they focus on better preparedness (factor 1) and expect spontaneous cohesiveness (harmony, factor 2). To ground our conclusion, we discussed it by referencing to existing evidence (see section 4.2.2 below).

4.2 Comparing to other studies

4.2.1 Absence of a leader

The present study found 114 out of 215 skiers (53%, 95%CI: 46%–60%) had their last backcountry ski trip in a group without informal or formal leader. Earlier field studies also show a considerable proportion of skiing trips were performed without a leader, e.g. 96/221 (43%, 95%CI: 37%–50%) (Aho-nen et al., 2024), and 8 out of 29 ski groups (28%, 95%CI: 13%–47%) (Zweifel and Haegeli, 2014).

Leadership is an extensively studied area, recognized as a crucial component of group dynamics, including in skiing as Zweifel (2015) highlighted. Burke et al. (2006) found that in teams with strong norms and interpersonal trust, members can function properly even in the absence of a leader. This is echoed by our finding that skiers with- or without-leader have consistently high levels of interpersonal trust with other group members and confidence in decision, even after controlling for individual skills.

4.2.2 Evaluative criteria of group performance

Few have studied if group performance is evaluated differently in the absence of a leader, e.g. among shared leadership (Zhu et al., 2018) and leader silence (Zill et al., 2020). Yet, these theories study a special form of leadership, rather than absence of a leader in a strict sense. For example, shared leadership can be the allocation of the leader role to members.

However, there is a wealth of research that demonstrates the critical role that processes such as communication and coordination play in team performance. LePine et al. (2008) did meta-analyses of relationships among teamwork processes, and found communication has positive relationships with team performance and member satisfaction. By evaluating simulated teams, Kilduff et al. (2000) showed that high-performing teams did not exhibit consensus at the beginning but converged as the team interacted. Our result echoed

their finding by showing skiers with- or without-leader always consider communication-related aspects in self-evaluating group performance (Openness and sense of involvement for with-leader, and harmony for without leader). Our results extend their research by uncovering that with no leader present, skiers have lower weight on communication in evaluating performance, and they seek for voluntary harmony rather than an active process of converge-to-consensus. Groups without an explicit leader are viewed to be at increased risk of negative group phenomena (Zweifel and Haegeli, 2014). For example, relying on spontaneous harmony risks falling into the dangerous trap of social proof heuristics (Couret et al., 2020) and groupthink (Taormina et al., 2014), where the desire for consensus overrides critical evaluation and leads to poor decision-making. As such, although this evaluative criterion in without-leader skiers is reasonable in terms of adapting to absence of leader, the effectiveness or correctness of it is questionable.

Furthermore, Mathieu et al. (2008) found that teams with a diverse skill set were more adaptable and capable of achieving high performance. The present study also found skiers with- or without-leader evaluate group performance using skill composition of group members (relevant competence for with-leader, and preparedness for without-leader). Notably, this is especially true for without-leader skiers, who place it as top consideration. This is consistent with existing evidence that highly skilled and experienced group members can substitute for leaders by providing peer guidance and support (Podsakoff et al., 1996). Also, skill redundancy can maintain performance levels even in the absence of leaders (LePine, 2005), suggesting that skiers with no leader count on the skill of their peers more. Adding to these findings, we further discovered a noteworthy difference between the type of skills skiers used to evaluate group performance between with- and without-leader skiers. Skiers with leader tend to focus on avalanche assessment skills, whereas skiers without leader consider a more diverse sets of skills, ranging from gear preparedness, to skiing skills, to avalanche assessment skills. Zweifel and Haegeli (2014) argued “The leader role is often rooted in personality traits, age, family structure or athletic skills, but not...avalanche expertise”. Three out of the five items under preparedness is not relevant to avalanche expertise (see table 3), which provides empirical support to this view. Notably, our results also suggest the phenomenon is less associated with groups with a leader.

Finally, we found skiers with a leader have another evaluative criterion, which is not present for those without a leader, that is sense of involvement (table 3, with-leader cohort, factor 3).

Table 3: Factor structure for evaluative criteria and definition

Items	Name	Nature/Definition
With-leader cohort		
factor 1 (VE: 23.1%)		
Everyone could voice their concerns to the leader (formal or informal) Everyone voiced their concerns whenever they felt necessary	Openness	This factor captures the degree to which group members feel free to express their opinions. It ensures all members' voices are heard and considered in decision-making.
factor 2 (VE: 21.3%)		
The least knowledgeable group member could conduct satisfactory avalanche assessments for this trip There was no large gap in avalanche assessment skills between the group members	Relevant competence	This captures the adequacy of avalanche assessment skills within the group, which ensures that all members can contribute effectively to safety assessments.
factor 3 (VE: 18.0%)		
The group decisions at the decision points were unanimous Everyone was happy with the decisions that were made	Sense of involvement	This is the sense of fulfillment everyone enjoys. It ensures all members feel their opinions are valued.
Without-leader cohort		
factor 1 (VE: 25.7%)		
The least knowledgeable group member could conduct satisfactory avalanche assessments for this trip There was no large gap in avalanche assessment skills between the group members There was no important difference in skiing skill level between group members, given the terrain All members were equipped with standard equipment and trained in the use of it The group members knew each other well	Preparedness	This factor reflects the overall readiness and capability of the group to handle avalanche terrain, which ensures that the group can operate effectively and safely in challenging conditions.
factor 2 (VE: 12.5%)		
Everyone in the group understood the decisions that were made Everyone was happy with the decisions that were made Decisions concerning avalanche hazard were well discussed in the group Everyone voiced their concerns whenever they felt necessary The group size was appropriate for the trip (time, difficulty) The roles of the group members were clearly defined	Harmony	Harmony ensures that the group operates smoothly, with mutual understanding, clear communication, and well-defined roles contributing to a collaborative and supportive environment.
Note: For English and Norwegian version of full items, go to https://osf.io/y2hxu ; VE: variance explained; Relevant competence: relevant competence of avalanche assessment		

4.3 Implications

Given the potential dangers of skiing in without-leader groups, and the lack of evidence that the altered evaluative criterion of prioritizing preparedness effectively mitigates risk, then in our view this criterion should be viewed as a psychological coping mechanism rather than reliable safety measure. We also recommend that avalanche courses should include the teaching of objective assessment capabilities for group quality, such as the SOCIAL checklist proposed by Zweifel (2015). This would help skiers make informed decisions when selecting groups to travel with. Additionally, our analysis provides support to develop the SOCIAL checklist further, by making explicit the latent criteria that skiers might want to think of when filling the checklist.

5. CONCLUSION

Compared to skiers with a leader, skiers travelling in avalanche terrain without a leader perceived equally acceptable interpersonal trust and confidence in decisions, and they even perceived a higher level of group dynamics. Nevertheless, skiers with and without a leader have different evaluative criteria for group performance, which match logically with leader presence/absence. The criteria for without-leader skiers, personal preparedness and expectation of cohesion, are arguably less sensible group-risk mitigation strategies as they may fail with the failure of one group member to comply.

References

- Ahonen, L., Landrø, M., Rong, G., Pfuhl, G., Mannberg, A., Hetland, A., and Cowley, B. U.: Decision Making in Avalanche Terrain: Insensitivity to Expert-level Evidence - A Recreationist Survey, *Journal of Outdoor Recreation and Tourism*, in press, 2024.
- Baumeister, R. F., Chesner, S. P., Senders, P. S., and Tice, D. M.: Who's in charge here? Group leaders do lend help in emergencies, *Personality and Social Psychology Bulletin*, 14, 17–22, 1988.
- Birkeland, K. W., Greene, E. M., and Logan, S.: In response to avalanche fatalities in the United States by Jekich et al, *Wilderness & Environmental Medicine*, 28, 380–382, 2017.
- Bryant, F. B. and Yarnold, P. R.: Principal-components analysis and exploratory and confirmatory factor analysis, American Psychological Association, Washington, DC, 1995.
- Burke, C. S., Stagl, K. C., Klein, C., Goodwin, G. F., Salas, E., and Halpin, S. M.: What type of leadership behaviors are functional in teams? A meta-analysis, *The leadership quarterly*, 17, 288–307, 2006.
- Carroll, J. M., Rosson, M. B., and Zhou, J.: Collective efficacy as a measure of community, in: Proceedings of the SIGCHI conference on human factors in computing systems, pp. 1–10, 2005.
- Casey-Campbell, M. and Martens, M. L.: Sticking it all together: A critical assessment of the group cohesion–performance literature, *International Journal of Management Reviews*, 11, 223–246, 2009.
- Castro, M., Bahli, B., Ferreira, J., and Figueiredo, R.: Comparing single-item and multi-item trust scales: Insights for assessing trust in project leaders, *Behavioral Sciences*, 13, 786, 2023.
- Cohen, J.: *Statistical Power Analysis for the Behavioral Sciences*, Routledge, New York, 2nd edn., 1988.
- Couret, R., Adam, C., and Mermillod, M.: Modelling risk-taking behaviour of avalanche accident victims, arXiv preprint arXiv:2009.09080, 2020.

- De Jong, B. A., Dirks, K. T., and Gillespie, N.: Trust and team performance: A meta-analysis of main effects, moderators, and covariates., *Journal of applied psychology*, 101, 1134–2016.
- Ebert, P. A. and Morreau, M.: Safety in numbers: how social choice theory can inform avalanche risk management, *Journal of Adventure Education and Outdoor Learning*, 23, 340–356, 2023.
- Ferguson, S. A. and LaChapelle, E. R.: *The ABCs of Avalanche Safety*, The Mountaineers Books, 2003.
- Galesic, M. and Bosnjak, M.: Effects of questionnaire length on participation and indicators of response quality in a web survey, *Public opinion quarterly*, 73, 349–360, 2009.
- Grosse, A. B., Grosse, C. A., Steinbach, L. S., Zimmermann, H., and Anderson, S.: Imaging findings of avalanche victims, *Skeletal radiology*, 36, 515–521, 2007.
- Hanafiah, M. H.: Formative vs. reflective measurement model: Guidelines for structural equation modeling research, *International Journal of Analysis and Applications*, 18, 876–889, 2020.
- Hendrikx, J. and Johnson, J.: Understanding global crowd sourcing data to examine travel behaviour in avalanche terrain, in: *Proceedings of the 2016 International Snow Science Workshop*, pp. 737–743, 2016.
- Kendall, D. L. and Salas, E.: Measuring team performance: Review of current methods and consideration of future needs, *The science and simulation of human performance*, pp. 307–326, 2004.
- Kerr, S. and Jermier, J. M.: Substitutes for leadership: Their meaning and measurement, *Organizational behavior and human performance*, 22, 375–403, 1978.
- Kilduff, M., Angelmar, R., and Mehra, A.: Top management-team diversity and firm performance: Examining the role of cognitions, *Organization science*, 11, 21–34, 2000.
- Kolb, J. A.: Leader behaviors affecting team performance: Similarities and differences between leader/member assessments, *The Journal of Business Communication*, 32, 233–248, 1995.
- LePine, J. A.: Adaptation of teams in response to unforeseen change: effects of goal difficulty and team composition in terms of cognitive ability and goal orientation., *Journal of Applied Psychology*, 90, 1153, 2005.
- LePine, J. A., Piccolo, R. F., Jackson, C. L., Mathieu, J. E., and Saul, J. R.: A meta-analysis of teamwork processes: Tests of a multidimensional model and relationships with team effectiveness criteria, *Personnel psychology*, 61, 273–307, 2008.
- Lewin, K.: Psychology and the process of group living, *The Journal of Social Psychology*, 17, 113–131, 1943.
- Light Shields, D. L., Gardner, D. E., Light Bredemeier, B. J., and Bostro, A.: The relationship between leadership behaviors and group cohesion in team sports, *the Journal of Psychology*, 131, 196–210, 1997.
- Mannberg, A., Hetland, A., Pfuhl, G., and Larsen, H. T.: Care panel, URL <https://uit.no/research/carepanel>.
- Mathieu, J., Maynard, M. T., Rapp, T., and Gilson, L.: Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future, *Journal of management*, 34, 410–476, 2008.
- McCammon, I.: Evidence of heuristic traps in recreational avalanche accidents, in: *Proceedings of the International Snow Science Workshop*, vol. 30, 2002.
- McCammon, I.: Heuristic traps in recreational avalanche accidents, *Work*, 406, 888–926, 2003.
- Page, C. E., Atkins, D., Shockley, L. W., and Yaron, M.: Avalanche deaths in the United States: a 45-year analysis, *Wilderness & Environmental Medicine*, 10, 146–151, 1999.
- Podsakoff, P. M., MacKenzie, S. B., and Bommer, W. H.: Transformational leader behaviors and substitutes for leadership as determinants of employee satisfaction, commitment, trust, and organizational citizenship behaviours, *Journal of management*, 22, 259–298, 1996.
- Procter, E., Strapazon, G., Dal Cappello, T., Castlunger, L., Staffler, H., and Brugger, H.: Adherence of backcountry winter recreationists to avalanche prevention and safety practices in northern Italy, *Scandinavian journal of medicine & science in sports*, 24, 823–829, 2014.
- Roberts, N. and Thatcher, J.: Conceptualizing and testing formative constructs: Tutorial and annotated example, *ACM sigmis database: The database for Advances in Information Systems*, 40, 9–39, 2009.
- Rosen, M. A. and Dietz, A. S.: Team performance measurement, *The Wiley Blackwell handbook of the psychology of team working and collaborative processes*, pp. 479–502, 2017.
- Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D. P., and Rafferty, L.: Is it really better to share? Distributed situation awareness and its implications for collaborative system design, *Theoretical Issues in Ergonomics Science*, 11, 58–83, 2010.
- Schmidt, J. A., Dunlop, P. D., and O'Neill, T. A.: Identifying the structure of within-team variance in ratings of team constructs, *Personnel Psychology*, 2023.
- Steiner, I. D.: *Group process and productivity*, Academic press New York, 1972.
- Taormina, R. J. et al.: Social harmony in decision making: Costs and benefits in Chinese society, *Open Journal of Social Sciences*, 2, 128, 2014.
- Tremper, B.: *Avalanche essentials: A step-by-step system for safety and survival*, Mountaineers Books, 2013.
- Tremper, B. and Diegel, P.: The wisdom of crowds in avalanche forecasting, in: *Proceedings of the International Snow Science Workshop*, Banff, Canada, pp. 78–84, 2014.
- Wallner, B., Moroder, L., Brandt, A., Mair, P., Erhart, S., Bachler, M., Putzer, G., Turner, R., Strapazon, G., Falk, M., et al.: Extrication times during avalanche companion rescue: a randomized single-blinded manikin study, *High Altitude Medicine & Biology*, 20, 245–250, 2019.
- Wongpakaran, T., Wongpakaran, N., Intachote-Sakamoto, R., and Boripuntakul, T.: The group cohesiveness scale (GCS) for psychiatric inpatients, *Perspectives in psychiatric care*, 49, 58–64, 2013.
- Zhu, J., Liao, Z., Yam, K. C., and Johnson, R. E.: Shared leadership: A state-of-the-art review and future research agenda, *Journal of Organizational Behavior*, 39, 834–852, 2018.
- Zill, A., Knoll, M., Cook, A., and Meyer, B.: When do followers compensate for leader silence? The motivating role of leader injustice, *Journal of Leadership & Organizational Studies*, 27, 65–79, 2020.
- Zweifel, B.: Group dynamics among recreationists traveling in avalanche terrain—challenges for improving avalanche safety, Ph.D. thesis, ETH Zurich, 2015.
- Zweifel, B. and Haegeli, P.: A qualitative analysis of group formation, leadership and decision making in recreation groups traveling in avalanche terrain, *Journal of Outdoor Recreation and Tourism*, 5, 17–26, 2014.
- Zweifel, B., Procter, E., Techel, F., Strapazon, G., and Boutellier, R.: Risk of avalanche involvement in winter backcountry recreation: the advantage of small groups, *Wilderness & environmental medicine*, 27, 203–210, 2016.