Integrating shear quality into stability test results

Ron F. Johnson[1] and Karl W. Birkeland[2,3]

[1]Gallatin National Forest Avalanche Center, Bozeman, Montana
[2]U.S. Forest Service National Avalanche Center, Bozeman, Montana
[3]Department of Earth Sciences, Montana State University

Abstract: This study investigates whether collecting shear quality data in conjunction with stability test results improves snowpack evaluations. Over the past six seasons we have consistently evaluated shear quality when evaluating snowpack stability. Shear quality is subjectively evaluated on a 3-tiered scale from Q1 (clean, fast shears) to Q2 (average shears) to Q3 (irregular or dirty shears). Our method is a formalization of what ski patrollers and others have been doing in the U.S. and elsewhere for at least several decades. We used a dataset of nearly 700 individual stability tests (rutschblock, stuffblock and compression tests) collected by seven observers on slopes from Alaska to Chile. In addition to stability test results, observers noted whether slopes they felt were similar to their snowpit location had avalanches, or collapsing or cracking snowpacks, on that day. Results suggest that shear quality provides important stability information, especially when stability test results appear to indicate relatively stable conditions, but the shear quality is rated Q1. This might be because stability test results are often spatially variable, while our experience indicates that shear quality is more homogeneous. Given these results, we believe formally integrating some description of shear characteristics into stability assessments may be important for avalanche workers and backcountry enthusiasts.

Keywords: avalanche forecasting, snow stability, stability tests, stability evaluation

1. Introduction

Avalanche workers and backcountry skiers use a variety of field stability tests to help assess avalanche danger. Most of the currently popular tests either quantitatively or qualitatively test the shear strength of the weak layer. Quantitative examples include the shear frame (Perla and Beck, 1983; Jamieson and Johnston, 2001) and the quantified loaded column test (Landry and others, 2001), while qualitative tests include the rutschblock (Fohn, 1987), compression (Jamieson and Johnston, 1997), and stuffblock (Birkeland and Johnson, 1999) tests. In addition to the actual test score, many avalanche workers informally evaluate the fracture character, fracture quality, and/or shear quality of the test. However, no consensus exists regarding the definitions or ratings of those attributes.

In a few cases, avalanche researchers and workers have begun to formally collect qualitative data on stability test fractures. In Canada, Jamieson (1995) and Jamieson and Johnston (2001) have observed twelve types of fracture surfaces in thousands of shear frame measurements. They found that only one type of fracture (when a divot greater than 10 mm existed under the rear compartment) exhibited significantly different strength values from the rest of the data. In addition, Parks Canada avalanche personnel use a system for evaluating what they term “fracture character” (Jamieson, 1999; Jamieson, pers. comm. 2002), and some recent work using this system will be presented at the 2002 International Snow Science Workshop (van Herwijnen and Jamieson, 2002). Their research defines fractures as Progressive Compression, Thin Planar, Sudden Collapse, or Non-Planar Breaks, and found that Thin Planar fractures comprised 70% of stability test fractures observed next to skier-triggered slabs in data from Canada’s Columbia Mountains. In Switzerland, Schweizer and others (1995) noted the importance of the type of release and the quality of the fracture when interpreting rutschblock tests. The type of release is described as “whole block”, “most of the block”, or “only a minor part of the block”, while quality of the fracture is rated as “clean”, “partly clean”, or “rough”. Schweizer and Weisinger (2001) discussed integrating this information into stability evaluations, and show that when avalanche forecasters rank the relative importance of several variables for interpreting rutschblock results, the highest ranked variables included type of failure and type of fracture plane, both of which ranked higher than the actual rutschblock score.

* *Corresponding author address:* Ron Johnson, Gallatin NF Avalanche Center, P.O. Box 130, Bozeman, MT 59771, USA; tel: 406-587-6984; fax: 406-587-6758; email: rjohnson@fs.fed.us

Table 1: Qualitative ratings of shear quality (from Birkeland and Johnson, 1999, *with significant additions in italics*).

Shear Quality	Description
Q1	Unusually clean, planar, smooth, *and fast* shear surface; weak layer may collapse during failure. Slab typically slides easily into the snow pit after weak layer fracture on slopes steeper than 35º, and sometimes on slopes as gentle as 25º. *Tests with thick, collapsible weak layers may exhibit a rougher shear surface due to erosion of basal layers as the upper block slides off, but the initial fracture was still planar and fast.*
Q2	“Average” shear; shear surface appears mostly smooth, but slab does not slide as readily as Q1. Shear surface may have some small irregularities, but not as irregular as Q3. Shear fracture occurs throughout the whole slab/weak layer interface being tested. The entire slab typically does not slide into snowpit.
Q3	Shear surface is non-planar, uneven, irregular, and rough. Shear fracture typically does not occur through the whole slab/weak layer interface being tested. After the weak layer fractures the slab moves little, or may not move at all, even on slopes steeper than 35º.

Independent of the work in Canada and Switzerland, we began to formally collect “shear quality” data in southwest Montana in the mid-1990s (Johnson and Birkeland, 1998; Birkeland and Johnson, 1999). Our method is based on what ski patrollers and other avalanche workers have been doing in the United States and elsewhere for at least the last 20 years and probably longer. Shear quality is subjectively evaluated on a 3-tiered scale from Q1 (clean, fast shears) to Q2 (average shears) to Q3 (irregular or dirty shears). As noted by Schweizer and Weisinger (2001), this is quite similar to the Swiss ratings for fracture quality. Our ‘shear quality’ actually mixes both the Swiss ‘fracture quality’ and ‘type of fracture’ into one rating, though our emphasis is on fracture quality. Our initial work did not adequately address the importance of shears fracturing on collapsible weak layers, so we updated our shear quality ratings to try to better reflect such weak layers (Table 1).

Despite the long history of avalanche workers noting shear quality, few researchers (with the exception of Schweizer and Weisinger (2001) and van Herwijnen and Jamieson (2002)) have rigorously tested the idea that these data are helpful for assessing snow stability. Further, though many experienced people feel this is important information, it is not formally included in the Canadian OGRS (CAA, 1995) nor is it typically taught in most avalanche classes or included in most avalanche books (one recent exception is Tremper, 2001). The purpose of this paper is to analyze almost 700 individual stability tests to see how useful our rating of ‘shear quality’ is for analyzing stability and how it might best be integrated into stability assessments.

2. Methods

At the beginning of the 2001/02 winter we asked a number of experienced avalanche workers who systematically note shear quality to share their data. Eight observers responded to our request, including backcountry avalanche forecasters, heli-skiing forecasters, backcountry guides, and avalanche educators. Observers recorded no result for 35 of 725 total tests. Due to the many factors that could lead to no result, we discarded those tests and utilized the other 690 tests, which included 149 rutschblock, 483 stuffblock, and 58 compression tests. Data compiled included the following: observer, date, mountain range, slope angle, test used, test score, shear quality (rated as Q1, Q2, or Q3), instability observed, avalanche danger for that day, and some snowpack information. This paper utilizes the stability test score, the shear quality, and the instabilities observed. The ‘instabilities observed’ category is somewhat subjective information that is important for this study. If an observer saw an avalanche, or noticed collapsing or cracking of the snowpack, in an area they felt was somewhat representative of the stability test location, then we considered instabilities to be present. Data were collected on slopes from Alaska to Chile, but the vast majority of the data are from the intermountain snow climates of Southwest Montana and Northwest Wyoming (Mock and Birkeland, 2000). Slope angles at the pit sites ranged from 24º to 45º, with an average angle of 32º.

First, we simply looked at the data for each of the three stability tests, comparing the percentage of tests associated with signs of instability with those not associated with instability for Q1, Q2, and Q3 shear qualities. After this initial analysis, we filtered the data to look at specific cases. We believe, as do others (i.e., Fesler, pers. comm.; McClung, 2000), that an avalanche danger assessment is essentially a search for instabilities. Therefore, we focused primarily on "false stable" cases, where stability test results indicate relatively stable conditions but signs of instability exist on similar slopes. We considered rutschblock scores of 5 or greater to indicate fairly stable conditions. Though some might argue that a rutschblock score of 5 does not indicate stable conditions, if we looked only at rutschblock scores of 6 we would have had insufficient data for our analyses. For stuffblocks, we analyzed drop heights of 0.40 m or greater. Previous work showed this is the median drop height associated with a rutschblock score of 5 (Birkeland and Johnson, 1999), and using a higher drop height made our data set too small. We had far fewer compression tests than either rutschblock or stuffblock tests, so we had to consider both "moderate" and "hard" results. We then compared the scores of each test, and the other signs of instabilities observed, for Q1, Q2, and Q3 shears. Due to the number of filters we applied to the data, the sample sizes are relatively small, and these results should be viewed with caution. However, we note sample sizes throughout, and our results roughly match our experience for utilizing shear quality data as a part of a stability assessment.

3. Results and discussion

Results suggest that our measure of 'shear quality' provides important stability information, especially when stability test results appear to indicate reasonably stable conditions, but the shear quality is rated Q1. For all rutschblocks (n=149), 65% of tests with Q1 shears were associated with signs of instability, dropping off to 28% of Q2 and 7% of Q3 shears (Figure 1). Similar percentages existed for the 58 rutschblock tests scoring 5 or 6, with 67% observations with Q1 shears (n=12), 30% of Q2 shears (n=33), and 0% of Q3 shears (n=13) being associated with signs of instability (Figure 2). Interestingly, when all 58 tests with scores of 5 or 6 are considered, 29% were associated with signs of instability and 71% were not, a result that mirrors that for the Q2 shears. These results clearly demonstrate what has been long understood; reasonably strong stability tests still can be associated with obvious signs of instability. However, these results suggest that those so-called "false stable" tests are much more likely to have a higher quality shear.

Somewhat similar results existed for stuffblock tests. For all tests (n=483), 49% of Q1, 34% of Q2 and 21% of Q3 shears were associated with signs of instability (Figure 3). A total of 96 tests had drop heights of 0.40 m or more, indicating relatively stable conditions. Thirty eight percent of those relatively stable tests that also had a Q1 shear were associated with other observed signs of instability (Figure 4). This percentage fell to 15% for both Q2 and Q3 shears. About 22% of the stuffblock tests with drop heights greater than or equal to 0.40 m were associated with other signs of instability, while 78% were not. Though not as dramatic as the rutschblocks, these stuffblock results provide further evidence that, for strong stability test results, Q1 shears are more commonly associated with signs of instability.

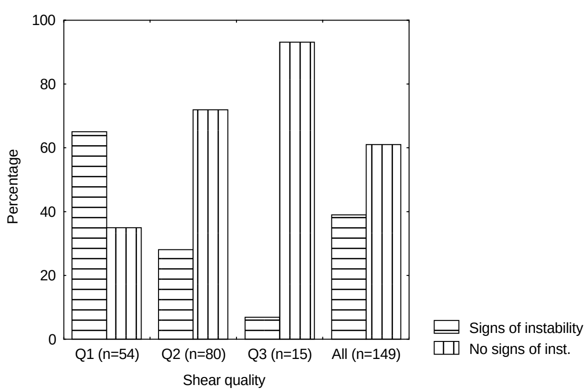

Figure 1: Shear quality and instabilities associated with all rutschblock tests (n=149).

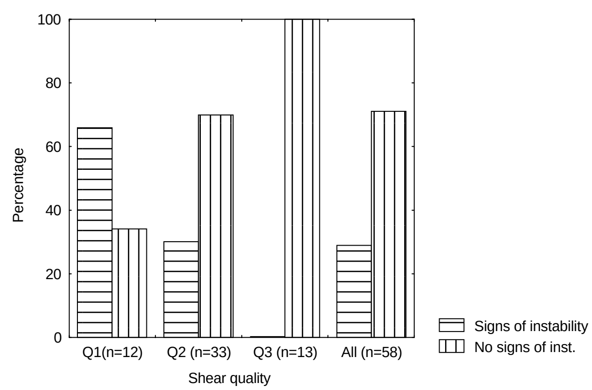

Figure 2: Shear quality and instabilities associated with all rutschblock scores of 5 or 6 (n=58).

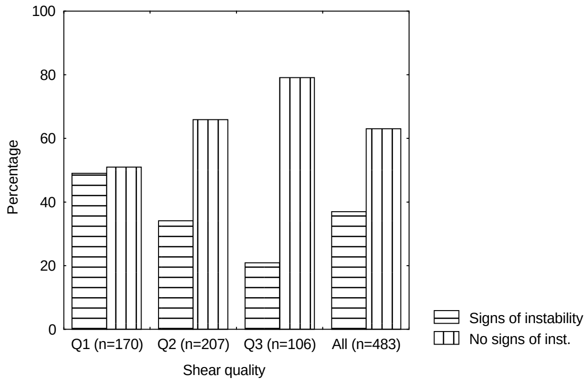

Figure 3: Shear quality and instabilities associated with all stuffblock tests (n=483).

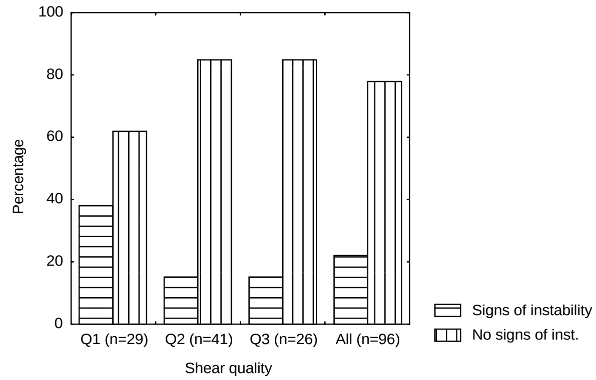

Figure 4: Shear quality and instabilities associated with stuffblock scores of 0.40 m or greater (n=96).

We had less data available for compression tests. Considering all 58 tests, 67% of Q1, 35% of Q2, and 18% of Q3 shears were associated with signs of instability (Figure 5). We had hoped to use only "hard" compression test results (CAA, 1995) for our analysis of "false-stable" results. However, we had only 23 tests in this category, which included no Q1 and only seven Q3 tests. Therefore, we expanded our analysis to include "moderate" test results, increasing the number of tests considered to 44. Using both "moderate" and "hard" test results, 100% of Q1 (n=5), 30% of Q2 (n=23), and 19% of Q3 (n=16) tests were associated with signs of instability (Figure 6). Again, a clear trend exists showing increasing percentages of higher quality shears being associated with observed evidence of instability.

Finally, the shear quality results from the aggregated dataset, for all tests types, suggest that Q2 shears do represent somewhat "average" conditions. Of the 690 tests, 35% were Q1, 45% were Q2 and 20% were Q3. Interestingly, for our data Q1 shears are much more common than Q3 shears. This could be the result of selective sampling by our observers, conducting their stability tests on days when higher-quality shears were more likely, or it may be because observers interpreted some Q3 shears to be the same as no result.

4. Conclusions

This research shows that deceptively strong stability test results with a Q1 shear quality are fairly commonly associated with other observable signs of instability. Therefore, integrating shear quality into stability test interpretations is important. It enhances the primary purpose of the stability test, which is a search for instability, reduces the uncertainty associated with a 'conditionally stable' stability test result, and helps to decrease the probability of making dangerous or possibly fatal errors (i.e., a "go" decision in a "no

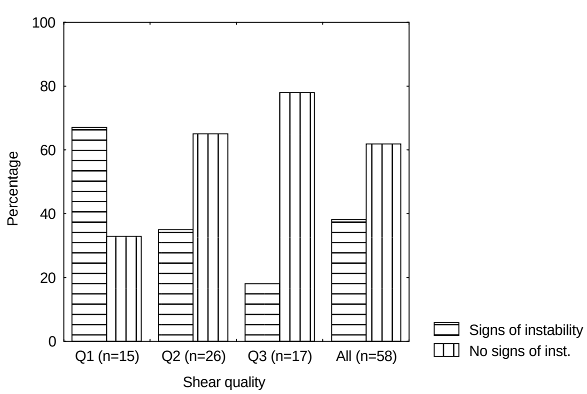

Figure 5: Shear quality and instabilities associated with all compression tests (n=58).

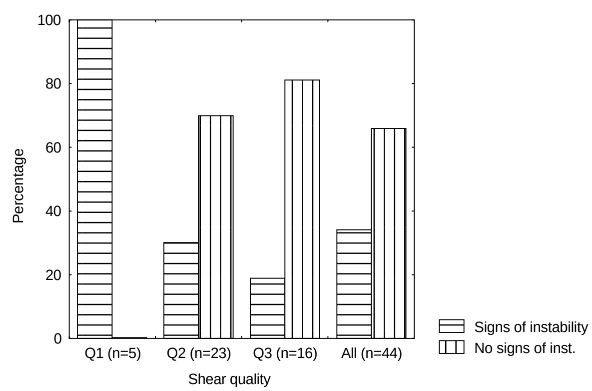

Figure 6: Shear quality and instabilities associated with moderate and hard compression tests (n=44).

go" situation). While the results of this study should be viewed with caution since the number of observations in some cases is still relatively small, we believe a formal integration of some measure of shear attributes will improve the interpretation of stability tests for avalanche workers as well as backcountry skiers, snowboarders and snowmobilers.

We do not know why Q1 shears are more commonly associated with instability. However, we believe shear quality provides important information about the relationship between the slab and the weak layer, which is a critical consideration for evaluating the avalanche potential (McClung and Schweizer, 1999). Perhaps shear quality provides a qualitative measure (at a small scale) of how well a fracture will propagate through the given weak layer.

We hypothesize at least two reasons why strong stability tests with high quality shears are more commonly associated with signs of instability than tests with lower quality shears. First, if a test is unknowingly conducted in a deeper part of the snowpack, the increase in effective depth might increase the amount of force needed to get the test to fail, even when the snowpack is unstable (Schweizer and Camponovo, 2001). Secondly, shear quality results may be more spatially homogeneous than the scores of the tests themselves. Landry (2002) documented high variability of shear strength on fairly uniform slopes utilizing the quantified loaded column test (QLCT) method (Landry et al., 2001). However, those trials suggested that shear quality in many cases was relatively homogeneous, and on a few sampling days all 50 tests conducted on a slope had Q1 shears even though shear strength and stability varied considerably. More variable shear qualities existed during tests on one day with a thick weak layer of older faceted snow. Thus, the spatially homogeneous shear quality results may be providing important information when a stability test happens to be unknowingly conducted in an area of the slope that has comparatively strong shear strength with respect to the weaker areas on the slope.

Given the observed relationships, we feel shear characteristics are one key component of stability analyses. However, the avalanche community has to decide the best method for documenting shear attributes. Currently there are at least three, and possibly more, methods. Some similarities exist between the Swiss method (Schweizer and Weisinger, 2001) and our 'shear quality' method, and the Parks Canada method (van Hervijnen and Jamieson, 2002) may complement these two systems. Hopefully the avalanche community can find a mutually agreeable method for formally noting a qualitative measure of shear attributes so that research efforts can be combined and results can be compared.

Finally, we emphasize that assessing 'shear quality' using any method is just one way to enhance the interpretation of stability test results, which are themselves only one component of a comprehensive assessment of the avalanche danger for a particular slope or area. Our data clearly show that even stability tests with high scores and a Q3 shear can be associated with readily observed signs of instability. As noted by Föhn (1987), interpreting stability tests requires experienced observers, and such tests must be supplemented with data such as snow profile evaluations, analyses of meteorological data, knowledge of recent avalanche activity, and knowledge of the terrain for a comprehensive, and holistic, evaluation of the avalanche danger.

Acknowledgements

Chris Landry reviewed this paper and provided valuable input and discussions. Our thanks go to those who collected the data: Doug Chabot, Scott Schmidt, Ian McCammon, Don Sharaf, Allan O'Bannon, Ethan Greene, and Tom Kimbrough. We are also grateful for the numerous conversations we've had on this subject over the past several years with many of our colleagues in the avalanche community.

References

Birkeland, K. and R. Johnson. 1999. The stuffblock snow stability test: comparability with the rutschblock, usefulness in different snow climates, and repeatability between observers. *Cold Reg. Sci. Tech.* 30, 115-123.

CAA, 1995. Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches. Canadian Avalanche Association, P.O. Box 2759, Revelstoke, BC, Canada, 98 pp.

Fesler, D. 1995. Personal communication. Avalanche educator and consultant, Anchorage, Alaska.

Föhn, P.M.B. 1987. The rutschblock as a practical tool for slope stability evaluation. *Int. Assoc. Hydro. Sci. Publ. 162*, 223-228.

Jamieson, B. 1999. The compression test – after 25 years. *The Avalanche Review*, 18(1), 10-12.

Jamieson, B. 1995. Avalanche prediction for persistent snow slabs. PhD dissertation, Dept. of Civil Eng., University of Calgary, Calgary, Alberta. 258 pp.

Jamieson, B. and C. Johnston. 1997. The compression test for snow stability. *Proc. of the 1996 Int. Snow Sci. Workshop*, Banff, Canada, 118-125.

Jamieson, B. and C. Johnston. 2001. Evaluation of the shear frame test for weak snowpack layers. *Ann. Glaciology* 32, 59-69.

Landry, C.C. 2002. Spatial variations in snow stability on uniform slopes: Implications for extrapolation to surrounding terrain. MS thesis, Dept. of Earth Sciences, Montana State University, Bozeman, Montana. 248 pp.

Landry, C.C., J.J. Borkowski, and R.L. Brown. 2001. Quantified loaded column stability test: mechanics, procedure, sample-size selection, and trials. *Cold Reg. Sci. Tech.* 33, 103-122.

McClung, D.M. 2000. Predictions in avalanche forecasting. *Ann. Glac.* 31, 377-381.

McClung, D. and J. Schweizer. 1999. Skier triggering, snow temperatures and the stability index for dry slab avalanche initiation. J. Glaciology, 45(150), 190-200.

Mock, C.J. and K.W. Birkeland. 2000. Snow avalanche climatology of the western United States Mountain Ranges. *Bulletin of the American Met. Soc.,* 81(10), 2367-2392.

Perla, R.I. and T.M.H. Beck. 1983. Experience with shear frames. *J. Glaciology* 29(103), 485-491.

Schweizer, J. and C. Camponovo. 2001. The skier's zone of influence in triggering slab avalanches. *Annals of Glaciology*, 32, 314-320.

Schweizer, J., C. Camponovo, C. Fierz, and P. Föhn. 1995. Skier triggered slab avalanche release – some practical implications. *Proc. Int. Symposium: Science and Mountain – The contribution of scientific research to snow, ice, and avalanche safety,* Chamonix, France, 30 May – 3 June. Association Nationale pour l'Etude de la Neige et des Avalanches (ANENA), Grenoble, France, p. 309-316.

Schweizer, J. and T. Weisinger. 2001. Snow profile interpretation for stability evaluation. *Cold Reg. Sci. Tech.* 33, 179-188.

Tremper, B. 2001. Staying alive in avalanche terrain. The Mountaineers, Seattle, Washington, 284 pp.

van Hervijnen, A. and B. Jamieson. 2002. Interpreting fracture character in stability tests. Abstract submitted to the 2002 International Snow Science Workshop, Penticton, British Columbia.