THE AVALUATOR'S OBVIOUS CLUES ACCIDENT PREVENTION VALUES: ARE THEY REPLICABLE?

Bob Uttl[1], Meaghen Henry[1,2], Jan Uttl[3]
[1]Red Deer College, Red Deer, AB, Canada
[2]University of Calgary, Calgary, AB, Canada
[3]Avidata.ca, Cochrane, AB, Canada

ABSTRACT: The Avaluator Avalanche Accident Prevention Card consists of the Trip Planner and Obvious Clues (Haegeli and McCammon, 2006). For Obvious Clues, the users count the number of obvious clues (e.g., avalanches, loading, terrain trap) and the Avaluator tells them the percentage of historical accidents prevented if historical victims had limited themselves to the same or fewer number of clues (i.e., relative risk reduction). However, the prevention values in the Avaluator differ widely from the values reported by the authors elsewhere as well as from various data sets reported by others. To illustrate, in response to our inquiries into these inconsistencies, Haegeli et al. (2006) recently (on April 16, 2008) revised the published prevention values from 90% to 98% for slopes with 2 or fewer clues and from 47% to 77% for 4 or fewer clues and explained the original values as “typos”. In light of these inconsistencies, we attempted to replicate the Avaluator's prevention values using several independent sets of avalanche accident reports. Across all independent data sets, our results reveal much lower prevention values than those reported in the Avaluator. In turn, our results strongly suggests that the prevention values reported in the Avaluator are biased and give users a false sense of security.

KEYWORDS: Accident Prevention, Risk, Decision Making

1. INTRODUCTION

In the 2002-03 winter season, 29 people died in avalanche accidents in Canada relative to the more common 10-15 avalanche fatalities each year. This unusual number of fatalities was caused principally by two accidents, each resulting in seven fatalities. One accident involved a school trip with final decision making authority held by two teachers with formal avalanche training whereas the other involved a commercial tour guided by several guides from Selkirk Mountain Experience (Penninman and Baumann, 2004). However, in terms of the number of avalanche accidents resulting in fatalities, a season with 14 fatal accidents was not at all remarkable (Uttl et al., 2008). Nevertheless, in response to the high number of fatalities, both Parks Canada (O'Gorman et al., 2003) and British Columbia government (Bhudak Consultants, 2003) commissioned reviews to identify possible improvements to public avalanche awareness and avalanche safety programs. The Parks Canada review (O'Gorman et al., 2003) recommended the development of “made-in-Canada” decision tools “to better equip recreational backcountry users to evaluate and reduce avalanche risk” (p. 70). O'Gorman et al. arrived at this recommendation in part based on their review of European avalanche accident risk reduction tools such as Munter's 3x3 method (Munter, 2002) as well as their consultations with Canadian experts, including Dr. Bruce Jamieson (University of Calgary) and Clair Israelson (Canadian Avalanche Centre). In response, the development of the Avaluator Avalanche Accident Prevention Card (Haegeli and McCammon, 2006) was sponsored by Parks Canada and funded by a New Initiative Fund grant from the National Search and Rescue Secretariat (NSRS), Canada, as part of the Canadian Avalanche Association's Decision Framework for Amateur Recreationists (ADFAR) project (Haegeli et al., 2006).

The Avaluator consists of the Trip Planner and Obvious Clues. The Trip Planner advises users on the suitability of a particular terrain for recreational activities given the current avalanche danger rating. The user obtains a terrain ATES rating (simple, challenging, complex) for an intended trip, as well as a current avalanche danger rating from the avalanche bulletin (low, moderate, considerable, high, extreme), and the Trip Planner makes one of the following recommendations: proceed with caution, proceed with extra caution, and not recommended.

The Obvious Clues (see Table 1) section advises users about a specific slope's stability; it is intended to help users “determine whether a slope is safe enough to cross” (Haegeli and McCammon, 2006, p. 14). The Obvious Clues is a checklist of seven clues to avalanche danger: danger rating of considerable or higher (Rating),

Corresponding author address: Bob Uttl, Red Deer College, 100 College Blvd, Red Deer, AB, T4N 5H5, Canada; email: uttlbob@gmail.com

avalanches in the area in the last 48 hours (Avalanches), signs of unstable snow (Unstable Snow), thaw instability (Thaw), loading by new snow, wind or rain in the last 48 hours (Loading), obvious avalanche path (Path), and terrain trap (Trap). The user adds up the number of Obvious Clues applicable to a specific slope and the Avaluator makes one of the three recommendations: proceed with normal caution (2 or fewer clues), proceed with extra caution (4 or 3 clues), and not recommended (5 or more clues). Moreover, the Avaluator shows users the percentage of historical accidents prevented if users had limited their travel to slopes with a given number of Obvious Clues. For example, the Avaluator states that if users had limited themselves to travel across slopes with 4 or fewer clues, 77% of historical accidents would have been prevented (p. 19, see Figure 1). Similarly, the Avaluator reports that limiting travel to 2 or fewer Obvious Clues would prevent 98% of accidents. While Haegeli and McCammon call this reduction in accidents "a prevention value", it is more commonly referred to as relative risk reduction.

Table 1. Avaluator's Obvious Clues (Haegeli and McCammon, 2006)

Avalanches: Are there signs of slab avalanche activity in the area within the last 48 hours?
Loading: Was there loading by snow, wind or rain in the area within the last 48 hours?
Path: Are you in an avalanche path or starting zone?
Terrain Trap: Are there gullies, trees or cliffs that would increase the consequences of being caught?
Rating: Is the danger rating considerable or higher?
Unstable Snow: Are there signs of unstable snow, such as whumpfing, cracking or hollow sounds?
Thaw Instability: Has there been recent significant melting of the snow surface by sun, rain or warm air?

However, our review of Haegeli, McCammon's and their colleagues writings about the Avaluator and the Obvious Clues method reveals a number of substantive inconsistencies, contradictory statements, errors, as well as failure to describe methods used to derive the prevention values in sufficient detail to allow their replication (Uttl et al., 2008). To illustrate:

- Haegeli and McCammon (2006; the Avaluator) claim that the prevention values reported in the Avaluator are based on North American accidents, Haegeli et al. (2006) claim that they are based on Canadian and US accidents, and McCammon and Haegeli (2007) claim that they are based on US accidents only. Thus, it is unclear whether the accidents are from US or from US and Canada.
- In the introduction to the Avaluator, Haegeli and McCammon (2006, p.1) state that they "studied over 1400 North American avalanche accidents" in developing the Avaluator, and in turn, this large sample of accidents confers a substantive sense of confidence in the prevention values based on such a large sample. However, the Obvious Clues prevention values reported in the Avaluator are actually based only on 252 accidents (McCammon and Haegeli, 2004, 2007). Thus, the Avaluator's prevention values are based on <18% of accidents remaining in their review after the authors deleted >82% of accidents and were left with a very selective and most certainly biased sample of avalanche accidents.
- While the deletion of >82% of accidents is not described in the Avaluator and elsewhere, McCammon and Haegeli (2004, 2007) note that they deleted all records with missing values, that is, where they could not establish the presence or absence of one or more obvious clues. Surprisingly, McCammon and Haegeli did not consider the impact that deletion of >82% of accidents would have on the prevention values (see Uttl et al., 2008, for a detailed discussion).
- The prevention values vary widely in different reports by the same authors. To illustrate, Figure 1 shows accident percentages as a function of the number of obvious clues reported by McCammon (2002, 2004) vs. Haegeli and McCammon (2006) in the Avaluator. Whereas McCammon's (2002, 2004) data imply prevention values of only 19.2 and 22.6%, respectively, for 4 or fewer Obvious Clues, the Avaluator data imply 77% prevention values. In turn, the Avaluator may be giving users a false sense of security and lead to more rather then fewer accidents.
- Even for *the sample of the same 252 accidents*, Haegeli et al. (2006) reported

different prevention values (90% for 2 or fewer clues, and 47% for 4 or fewer clues) than the values reported in the Avaluator (98% and 77%). In response to our inquiries, on April 16, 2008, Haegeli corrected the values in Haegeli et al. (2006) to be consistent with the Avaluator. However, Haegeli did not respond to our inquiries whether he re-analyzed the data and made sure that the “typos” were not in the Avaluator as opposed to in Haegeli et al. (2006).

- Another critical step in analyses of accident records – assessment of coding reliability and coder bias – is not mentioned anywhere in Haegeli, McCammon and their colleagues' writings. Moreover, the analysis of discrepancies between McCammon (2000) and McCammon (2004) data suggest very poor coding reliability (see Uttl et al., 2008, for details).

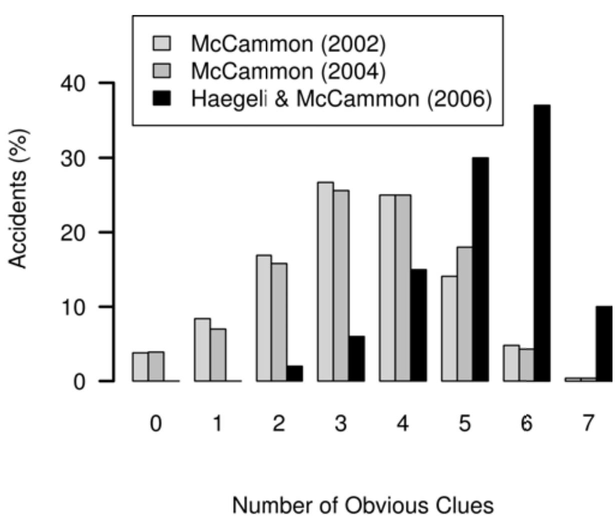

Figure 1. Accident percentage as a function of the number of obvious clues reported by McCammon (2002, 2004) vs. Haegeli and McCammon (2006) in the Avaluator.

Moreover, despite repeated requests, we were unable to obtain from the authors (Haegeli and McCammon) methodological clarifications, the list of accidents included in the 252 selective sample, and access to the raw data used to derive the prevention values reported in the Avaluator, even though we requested them only for the limited purpose of verifying the authors' substantive claims.

Indeed, Uttl et al. (2008) review of the literature surrounding the Avaluator and Obvious Clues method have shown that the most likely candidate for the large discrepancy between McCammon (2002, 2004) vs. the Avaluator data is the treatment of missing values. McCammon (2002, 2004) did not exclude missing values and implicitly assumed that missing values mean that Obvious Clues were not present. In contrast, Haegeli and McCammon (2006, 2007) deleted cases with missing values and implicitly assumed that the remaining 18% of accidents (252 accidents) were still a representative and an unbiased sample of all available records of avalanche accidents (itself a subsample of all avalanche accidents as many accidents go unreported).

However, this later assumption that the missing values were missing completely at random (MCAR) is rarely tenable (Little and Rubin, 1987; Schaffer and Graham, 2002). Victims, witnesses, rescuers, and investigators are likely to report the presence of the Obvious Clues precisely because they are “obvious”, apparent to anyone on site (Baddeley, 1990). In contrast, evidence from cognitive science, memory research, as well as eyewitness research studies indicates that people rarely mention the absence of things (Baddeley, 1990; Brewer and Treyens, 1981; Uttl and Graf, 2006), and thus, in the case of Obvious Clues, the missing values probably mean that specific Obvious Clues were not present. In turn, excluding accidents with missing values is inappropriate because it excludes accidents with fewer clues, biases the final sample towards accidents with many Obvious Clues, and renders the final sample non-representative.

We attempted to replicate the prevention values published in the Avaluator using the sample of avalanche accidents compiled in *The Snowy Torrents Avalanche Accidents in the United States, 1980-1986* (Logan and Atkins, 1996). This set of accidents was also included in the Avaluator's initial sample (i.e., prior to accident deletion due to missing values). In contrast to Haegeli and McCammon (2006) and McCammon and Haegeli (2007) we did not simply delete accidents with missing values but calculated the frequency distribution of the Obvious Clues and associated prevention values under different assumptions about missing values to establish the lower and upper bounds of prevention values in light of the uncertainty due to missing values (Little and Rubin, 1987; Schaffer and Graham, 2002).

Next, we used external data sources to obtain external estimates of Obvious Clue prevalence to establish the likely prevalence of Obvious Clues and likely prevention values within the regions of uncertainty caused by missing values. Specifically, we used National Oceanic and Atmospheric Administration (NOAA) Global Summary of Day (SOD) Surface Data to estimate Thaw prevalence and we used Avalanche Bulletins to estimate the prevalence of Thaw and Unstable Snow in Canada. Whereas NOAA weather data provide objective estimates of Thaw, Avalanche Bulletins provide estimates of Rating, Thaw and Unstable Snow based on subjective records issued by specific avalanche professionals. However, Avalanche Bulletins Danger Ratings have only moderate agreement with the local danger ratings, with agreement ranging from 57% to 64% (Jamieson et al., 2006).

In turn, external estimates allow us to determine what missing values really mean: Do they mean the absence of Obvious Clues as suggested by cognitive science and memory research? Or are missing values merely missing due to some random process as unwittingly assumed by Haegeli and McCammon for the Avaluator?

Previous research suggests that it is far more likely that victims, witnesses, rescuers, and investigators did not mention specific Obvious Clues because they were not present in specific accidents with one exception: Avalanche Danger Rating. The Avalanche Danger Rating is not directly observable by victims, witnesses, rescuers, or investigators on site but must be sought out, retrieved from websites, telephone information services, public bulletins, ranger stations, information offices, and the like. Moreover, the Avalanche Danger Rating is often unavailable for specific accidents because it was historically not even issued. Thus, if the avalanche danger rating was not provided in the record we assessed the avalanche danger rating ourselves using the standard avalanche danger rating scale: low, moderate, considerable, high, and extreme.

2. METHOD

2.1 *Accident Sample*

We reviewed all 146 accidents in *The Snowy Torrents Avalanche Accidents in the United States, 1980-1986* (Logan and Atkins, 1996). Each accident report contains information provided by the survivors, witnesses, rescuers, and other individuals involved as well as comments by the authors. Many of the reports include drawings of accident sites with positions of individuals involved in the accidents.

For our analysis, we excluded highway accidents, accidents involving structures (e.g., buildings) rather than trips into avalanche terrain, and one incident where no people were involved. We did not exclude accidents involving commercial guides as well as other professionals (e.g., snow rangers), because avalanches do not take the professional status of their victims into account, and because our preliminary findings show that commercial guides and other professionals make similar and perhaps even more severe decision making errors than recreationalists (e.g., Selkirk Mountain Experience Durrand Glacier incident, Penninman and Baumann, 2004). Thus, in total, our analysis is based on 131 accidents involving trips into terrain capable of producing avalanches.

2.2 *Obvious Clue Coding and Coding Reliability*

For each accident we coded basic descriptive information including the date, time, location, and activity.

Critically we coded the presence of Obvious Clues using the Avaluator (Haegeli and McCammon, 2006) (see Table 1). However, because the presence or absence of the Obvious Clues in many records as well as in real life is less than “obvious” (think about deciding whether an avalanche ran within 48 hours or earlier), we coded the presence of each Obvious Clue using the following scale: Yes (the record shows the clue was present), Weak Yes (it is likely the clue was present although not certain), Don't know (the record does not allow us to determine the presence or absence of the clue), Weak No (it is likely the clue was absent though not certain), and No (the record shows the clue was not present).

To establish inter-rater agreement and reliability, a randomly selected 41 accidents were independently coded by a second coder and inter-rater agreement and reliability was calculated for the coding of each Obvious Clue as well as the sum of Obvious Clues.

Finally, whenever the Avalanche Danger Rating (the only clue not directly observable by victims, witnesses, rescuers and investigators on site) was unknown for a specific accidents, the coder assigned an avalanche danger rating using the standard scale ranging from low to extreme. If the information was insufficient to assign the rating, the missing value code was assigned.

2.3 External Estimates of Obvious Clues

To obtain external estimates of Thaw for each accident in our sample, we determined the longitude and latitude of each accident from the entire accident record and using automated search retrieved accident specific temperature data for the nearest National Oceanic and Atmospheric Administration (NOAA) weather station with Global Surface Summary of Day (SOD) Data available for the day of the accident. The retrieved temperature data were then corrected for the elevation difference between the weather station and the accident location elevation. If a day's maximum elevation adjusted temperature climbed over 0 Celsius (freezing), we coded Thaw as possibly present and otherwise as absent. We also obtained external estimates of Thaw for accidents reported in *Avalanche Accidents in Canada Volume 4 1984-1996* (Jamieson and Geldsetzer, 1996) using the same methodology.

To obtain external estimates for Ratings, Unstable Snow and Thaw Instability we also used Avalanche Bulletins associated with Canadian accidents archived by Cyberspace Avalanche Center (www.csac.org) from October 1, 2000 to May 31, 2008. To establish an estimate of signs of Unstable Snow, we searched the bulletins for occurrences of signs of unstable snow including whumpfing, cracking, hollow sounds, and collapsing. To establish an estimate of signs of Thaw, we searched the bulletins for signs of Thaw including wet snow, moist snow, mushy snow, and rain.

3. RESULTS

3.1 Preliminaries

As noted above, *Snowy Torrents Avalanche Accidents in the United States, 1980-1986* (Logan & Atkins, 1996) includes 146 avalanche accident reports. However, after excluding accidents involving highway traffic and work, permanent structures (e.g., buildings), and incidents involving no people (1 report), we retained 131 reports for the analysis.

3.2 Coding Reliability

To assess inter-rater agreement and coding reliability, a second coder independently coded 41 randomly selected accidents. For the "Yes" (clue was present) only criterion, the agreement percentages were: Avalanches 93%, Loading 85%, Path 83%, Traps 81%, Rating 93%, Unstable Snow 95%, and Thaw 90%. Similarly, for the "Yes" and "Weak Yes" (the clue was present or probably present) criterion, the percentage agreements were: Avalanches 95%, Loading 88%, Path 95%, Traps 81%, Rating 95%, Unstable Snow 90%, and Thaw 88%.

Figure 2 shows the prevalence of individual cues using "Yes" and "Weak Yes" criteria for the two coders. This figure highlights high coder agreement and absence of bias. Correlation between the mean clue prevalence determined by the two coders was 0.99 (n = 7).

Inter-rater agreement calculated for the number of the Obvious Clues (i.e., the Avaluator Obvious Clue score) using "Yes" and "Weak Yes" criteria was respectable but not perfect: the two coders detected the same number of Obvious Clues in 51% of cases, differed by 1 clue in 39% of cases, and differed by 2 clues in the remaining 10% of cases. Correlation between the number of Obvious Clues detected by the two coders was 0.67 (n = 41), respectable for a 7-item checklist but far from perfect. All things being equal, longer checklists such as 3.57 times longer Bolognesi 25 item Nivo Test (2007) are more reliable (Spearman-Brown Prophecy formula yields reliability 0.88 for 25 item vs. 7 item checklist) (Crocker and Algina, 1986) as well as more tolerant of partial information input or user input errors (Bolognesi, 2004, quoted in McCammon and Haegeli, 2005).

Finally, the average number of clues detected by coder 1 and 2 was comparable: 2.80 vs. 2.85.

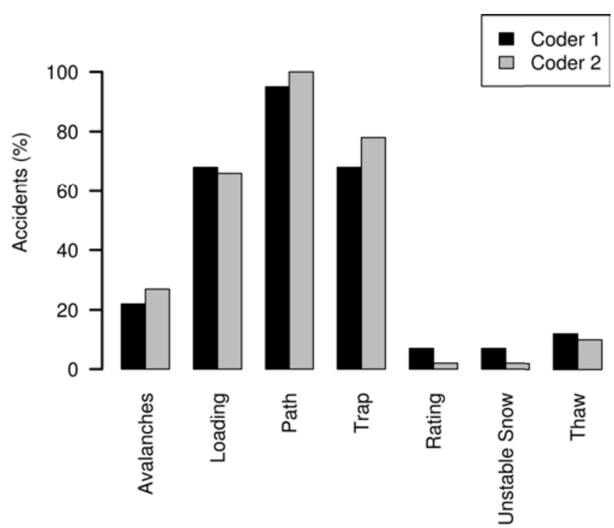

Figure 2. Prevalence of Individual Obvious Clues using "Yes" and "Weak Yes" criteria for the two coders (over 41 randomly selected accidents).

3.3 *Prevalence of Obvious Clues from Accident Records and External Data*

The stacked bars in Figure 3 show the prevalence of Yes (clue is present), Weak Yes (clue is probably present), Unknown/DNK (presence or absence of clue cannot be established), Weak No (clue is probably absent), and No (clue is absent) judgments for each Obvious Clue. As noted above, in contrast to all other clues, the Rating clue is not observable at the accident site but must be looked up by victims, witnesses, rescuers or investigators. Thus, for the Rating clue, Weak Yes and Weak No represent our assignment of Rating as Considerable or higher and Moderate and lower, respectively, based on information provided in the accident records. The DNK represents the proportion of accidents with insufficiently detailed information to allow us to assign an avalanche danger rating.

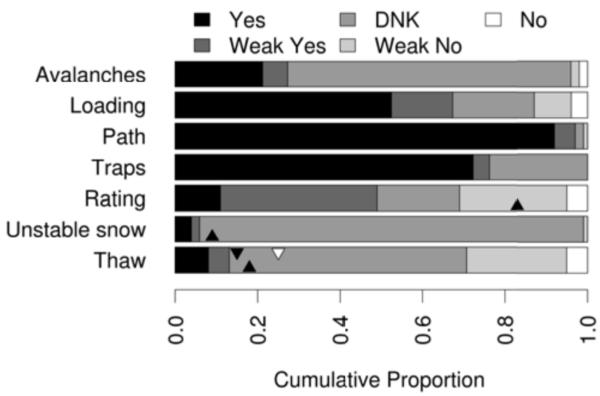

Figure 3. Prevalence of Obvious Clues Judgments: Yes (present), Weak Yes (probably present), DNK (do not know), Weak No (probably absent), No (Absent). Upwards pointing triangles show external estimates of Thaw, Unstable Snow, and Rating obtained from Canadian Avalanche Bulletins whereas downwards pointing triangles shows external estimate of Thaw obtained from NOAA weather station data for the current sample of accidents (Snowy Torrents, white triangle) and for Canadian accidents (black triangle). There are large proportions of accidents where the presence or absence of each Obvious Clue is indeterminate from the records (DNK judgments, middle gray). Moreover, the external estimates of Obvious Clues are strong evidence that missing information about Obvious Clues means that clues were absent in the vast majority of cases.

The DNK portions (middle gray) in Figure 3 highlight that accident records themselves do not allow us to determine the presence or absence of Obvious Clues for large portions of accidents. To illustrate, signs of Avalanche activity within 48 hours were described in 21% of records, and signs of Avalanche activity that probably occurred within 48 hours but not necessarily were present in another 6% of records. Only 2% of records noted absence of signs of Avalanche activity within 48 hours and another 2% allowed the conclusion that signs of Avalanche activity within 48 hours were probably absent. The remaining 68% of records (DNK portion) allow no conclusion about the presence or absence of Avalanche clue from the records themselves.

In general, cognitive science, memory, and eyewitness research indicates that victims, witnesses, rescuers, and investigators are unlikely to report the absence of things, and therefore, missing values (DNK portion) most likely means that a given clue was absent. This conclusion is supported by external estimates of Obvious Clues.

The downwards pointing triangles in Figure 3 show external estimates of Thaw obtained from objective weather data (NOAA weather stations closest to accident sites) for the current accident sample (Snowy Torrents) and for Canadian accidents described in Jamieson and Geldsetzer (1996). In contrasts, upwards pointing triangles show external estimates of Thaw, Unstable Snow, and Rating obtained from Avalanche Bulletins for the most recent Canadian accidents (October 2001 to May 2008) archived by Cyberspace Snow and Avalanche Center (www.csac.org). The external estimates of Obvious Clues for Thaw and Unstable Snow clues indicate that whenever these clues are not mentioned in records, they are very likely absent. In turn, this evidence suggests that whenever observable clues (Avalanches, Loading, Path, Traps, Unstable Snow, Thaw) are not mentioned in records, they were highly likely absent.

In contrast, the missingness mechanism for a single non-observable clue – Rating – is certainly different than for observable clues. First, the Rating must be assigned and we know that Ratings were not issued for many accident sites described in Snowy Torrents (Logan and Atkins, 1996). Second, the Rating must be looked up rather than observed by victims, witnesses, rescuers, or investigators on site. Thus, it is not surprising that external estimate of Rating obtained from Canadian data is slightly higher than the Rating obtained from the records (Yes) and supplemented by the coder's assessment of avalanche danger rating using information available from records (Weak Yes).

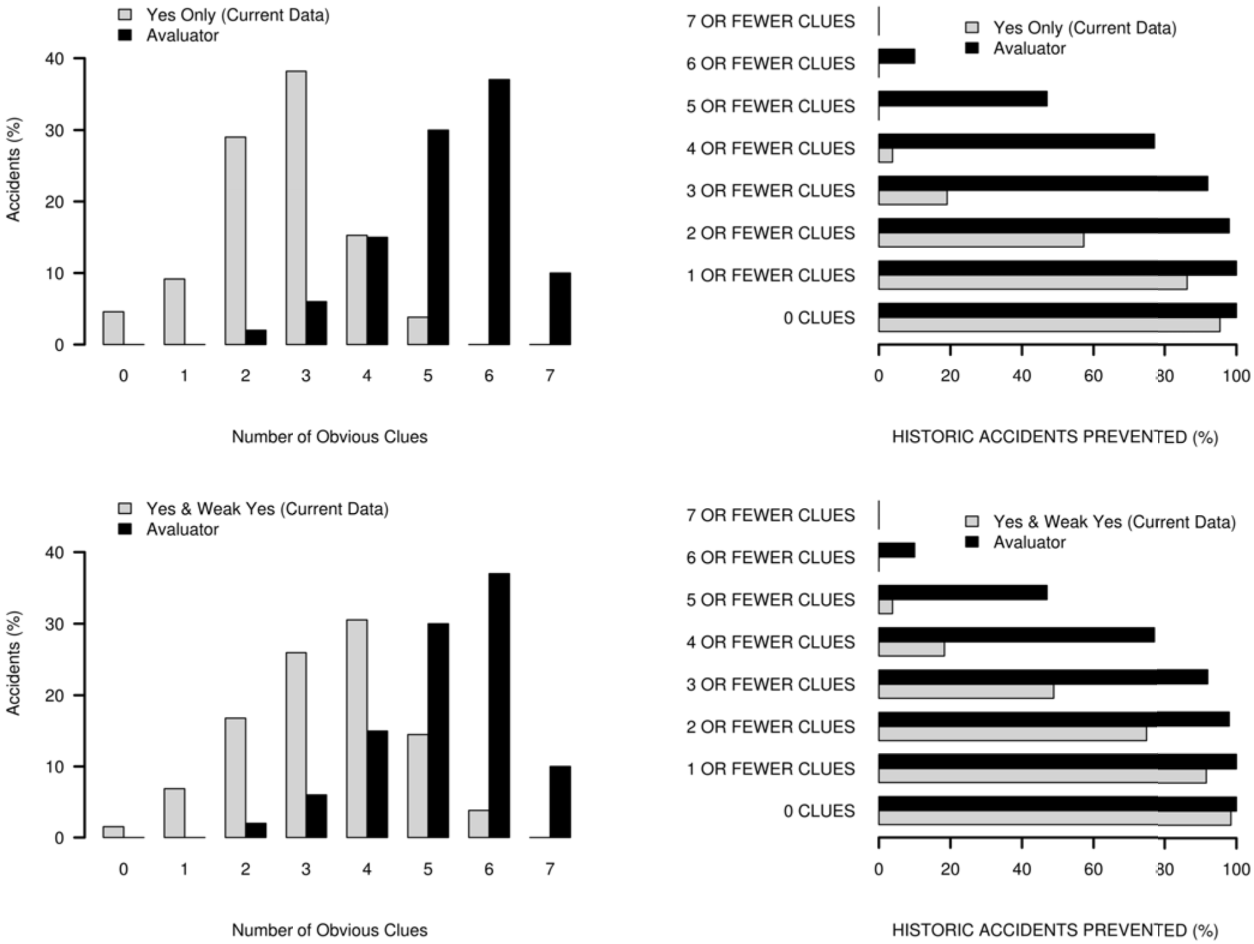

Figure 4. Panel A (top left) shows the distribution of Obvious Clues for current data using "Yes" clue was present criteria only (gray bars) and in the Avaluator (black bars). Panel B (top right) shows the corresponding prevention values (i.e., relative risk reduction values) for current data using "Yes" clue was present criteria (grey bars) and the Avaluator data (black bars). Our data yield a prevention value of only 4%, whereas the Avaluator claims a prevention value of 77%. Panels C (bottom left) and D (bottom right) show the distribution of Obvious Clues and the prevention values for current data using both "Yes" and "Weak Yes" criteria (i.e., clue was present or probably present) and the Avaluator. Similar to the "Yes" clue is present criterion, the "Yes" and "Weak Yes" criteria yield much smaller prevention values than the Avaluator data: 18% vs. 77%, respectively.

3.4 *Sensitivity Analysis: Distribution of Obvious Clues and Prevention Values under Different Assumptions*

Figure 4 Panel A (top left) shows the distribution of Obvious Clues for current data using the "Yes" clue was present criteria only (gray bars) and in the Avaluator (black bars). Panel B (top right) shows the corresponding prevention values (i.e., relative risk reduction values) for current data using the "Yes" clue was present criteria (grey bars) and the Avaluator data (black bars). Our data yield a prevention value of only 4% whereas the Avaluator claims a prevention value of 77%.

The results from more liberal criteria for the presence of Obvious Clues, "Yes" and "Weak Yes" judgments, are shown in Figure 4 Panels C (bottom left) and D (bottom right). Similar to "Yes" clues is present criteria, "Yes" and "Weak Yes" criteria yield much smaller prevention values than the Avaluator data: 18% vs. 77%, respectively.

The distribution shown in Figure 4 Panel C is very similar to the distribution reported by McCammon (2002, 2004) (see Figure 1 above) and yields similar prevention values for 4 or fewer

Obvious Clues. Our “Yes” and “Weak Yes” criterion data yield a prevention value of 18% whereas McCammon's data yield prevention values of 19.2% (2002) and 22.6% (2004). In contrast, the Avaluator distribution of Obvious Clues is shifted markedly towards a higher number of clues and, therefore, higher prevention values.

However, external estimates of Obvious Clues are in agreement with the “Yes” and “Weak Yes” criteria and strongly support the assumption that if an observable clue is not mentioned in a record, it was most likely not present. In turn, external estimates highlight that true distributions and true prevention values are close to our current data gathered under the assumption that missing values for observable obvious clues indicate the absence of obvious clues.

In contrast, external estimates of Obvious Clues highlight that the assumption that missing values in avalanche records are missing completely at random (MCAR) is false. In turn, the Avaluator's prevention values are biased, indeed impossible in light of the external estimates of just three of the seven Obvious Clues. The Avaluator data indicate that historical avalanche accidents had on the average 5.24 Obvious Clues present. However, for Canadian data, independent estimates of only three of the Obvious Clues (Thaw: 0.15; Unstable Snow: 0.09; Rating: 0.84) reduce the maximum average number of the Obvious Clues to 5.08 (from a theoretical maximum of 7), well below the 5.24 average for the Avaluator.

4. DISCUSSION

Our independent analysis of avalanche accident records published in *Snowy Torrents Avalanche Accidents in the United States, 1980-1986* (Logan & Atkins, 1996) demonstrates that the presence or absence of Obvious Clues is indeterminate from the records in a large proportion of accidents, resulting in a large degree of uncertainty about the status of each Obvious Clue. To increase confidence in our results, we have also conducted independent analyses of two more accident samples including accident records published in *Avalanche Accidents in Canada Volume 4 1984-1996* (Jamieson and Geldsetzer, 1996) and obtained substantially the same results.

In turn the question arises what do the missing values mean? While exceptions exist, from cognitive science, memory, and eyewitness research we know that victims and witnesses rarely mention the absence of things (Baddeley, 1990; Brewer and Treyens, 1981; Uttl and Graf, 2006), indicating that whenever an Obvious Clue is not mentioned in an accident record it probably means that the Obvious Clue was not present.

Moreover, objective external estimates of the Thaw clue prevalence from the NOAA weather records as well as subjective external estimates of Thaw and Unstable Snow clues obtained from Canadian avalanche bulletins are in close agreement with the prevalence of the clues derived from accident records, assuming that the missing values (i.e, when the status of a clue is indeterminate or DNK judgments) mean absence of the clues.

In turn, the objective and subjective external estimates of Obvious Clues highlight that the list-wise deletion of > 82% of accident records with missing values (i.e, values indicating absence of clues) adopted by the Avaluator authors resulted in a severely biased sample of accident records and consequently, severely biased prevention values reported in the Avaluator.

Figure 4 shows that the degree of bias is huge with deadly consequences. Using the most conservative criteria for deciding whether a clue is present or absent (Yes judgments only), users who limit their travel to 4 or fewer clues will avoid only 4% of accidents. Using more liberal criteria (Yes and Weak Yes judgments) that are consistent with external objective and subjective estimates of Obvious Clues, users who limit themselves to 4 or fewer clues will avoid only 18% of accidents. In contrast, the Avaluator prevention values based on <18% of accident records selected through inappropriate listwise deletion of cases with missing values claims a prevention value of 77% for 4 or fewer clues.

In summary, the Avaluator prevention values are not replicable; they are the result of inappropriately deleting > 82% of avalanche accidents and other methodological artifacts such as unreliable coding of Obvious Clues (see Uttl et al., 2008). The Avaluator gives users a false sense of confidence in the stability of slopes they are just about to cross. In turn, use of the Avaluator will lead to more avalanche accidents and more avalanche accidents deaths.

5. CONCLUSIONS

The Avaluator was developed by Haegeli and McCammon (2006) as part of the Canadian Avalanche Association ADFAR project (funded in excess of $600,000). It was released on the market in October 2006 and within the 2006-07 season over 7,500 Avaluators were sold. Moreover, the Avaluator has been incorporated into Canada's public avalanche awareness and prevention programs and is now part of the curriculum of Avalanche Safety Training courses. The aim of a recently launched successor project, ADFAR2 ($559,030.60 funded from New Initiative Funds of Search and Rescue Secretariat, Canada), is “a detailed evaluation of the Avaluator's efficacy” (Haegeli, www.avisualanche.ca, retrieved July 22, 2008). According to NSRS website, Dr. Pascal Haegeli is the Principal Investigator charged with the evaluation of the Avaluator (Haegeli and McCammon, 2006). Thus, Haegeli who is the principal author of the Avaluator, and who (together with Ian McCammon) has substantive commercial interests in the success of the Avaluator, has been charged with the evaluation of the Avaluator's efficacy.

Moreover, Haegeli and McCammon have repeatedly refused to clarify their method and to provide access to their data even for the limited purpose of verification of their claims when we raised concerns about numerous inconsistencies, contradictory statements, and errors in their writings about the Avaluator and Obvious Clues methods (Uttl et al., 2008).

Thus, the exact method and scientific basis of the Avaluator, if any, is held in tight secrecy and is not available for either scientific or public inspection. However, our independent analysis shows that the Avaluator will facilitate rather than prevent avalanche accidents by conferring a false sense of security about slope stability to its users.

6. ACKNOWLEDGMENTS

We thank Mekale Kibreab, Kelly Kisinger, and Dylan Smibert for help with data collection and Amy L. Siegenthaler for careful reading and comments on the manuscript.

7. REFERENCES

Baddeley, A., 1990. Human Memory. London: Allyn and Bacon.

Bhudak Consultants, 2003. Public Avalanche Safety Program Review. Report prepared for the British Columbia Ministry of Public Safety and Solicitor General.

Bolognesi, R., 2007. Avalanche! Understand and reduce the risks from avalanches. Cumbria, UK: Cicerone.

Brewer, W. F., & Treyens, J. C., 1981. Role of Schemata in Memory for Places. Cognitive Psychology, 13, 207-330.

Crocker, L., & Algina, J., 1986. Introduction to Classical & Modern Test Theory. New York: Harcourt Brace Jovanovich.

Haegeli, P., & McCammon, I., 2006. Avaluator Avalanche Accident Prevention Card. Revelstoke, BC: Canadian Avalanche Association.

Haegeli, P., McCammon, I., Jamieson, B., Israelson, C., & Statham, G., 2006. The Avaluator – A Canadian rule-based avalanche decision support tool for amateur recreationists. Proceedings of the International Science Workshop, Telluride, CO.

Jamieson, B., & Campbell, C., & Jones, A., 2006, October. Spatial and Time Scale Effects in Canadian Avalanche Bulletins. Proceedings of the International Snow Science Workshop, Telluride, CO.

Jamieson, B., & Geldsetzer, T., 1996. Avalanche Accidents in Canada Volume 4 1984-1996. Revelstoke, BC: Canadian Avalanche Association.

Little, R. J. A., & Rubin, D. B., 1987. Statistical analysis with missing data. New York: Wiley.

Logan, N., & Atkins, D., 1996. The Snowy Torrents Avalanche Accidents in the United States 1980-1986. Denver, CO: Colorado Geological Survey.

McCammon, I., 2002. Evidence of heuristic traps in recreational avalanche accidents. International Snow Science Workshop, Penticton, BC.

McCammon, I., 2004. Heuristic traps in recreational avalanche accidents: Evidence and implications. Avalanche News, 68, 1-11.

McCammon, I., & Hageli, P., 2004. Comparing avalanche decision frameworks using accident data from the United States. International Snow Science Workshop, Jackson, WY.

McCammon, I., & Haegeli, P., 2005. Description and evaluation of existing European decision-making support schemes for recreational backcountry travelers. Report for Canadian Avalanche Association, NIF Project: Avalanche Decision Framework for Amateur Winter Recreationists.

McCammon, I., & Haegeli, P., 2007. An evaluation of rule-based decision tools for travel in avalanche terrain. Cold Regions Science and Technology, 47, 193-206.

Munter, W., 2002. Drei mal drei (3x3) Lawinen. Bergverlag Rother.

O'Gorman, D., 2003. Parks Canada's Backcountry Avalanche Risk Review. Report of the Independent Panel prepared for Parks Canada.

Penniman, D., & Baumann, F., 2004. The SME Avalanche Tragedy of January 20, 2003: A Summary of the Data. International Snow Science Workshop, Jackson Hole, WY.

Schaffer, J. L, & Graham, J. W., 2002. Missing data: Our view of the state of the art. Psychological Research, 7, 147-177.

Uttl, B., & Graf, P., 2006. Age-Related Changes in the Encoding and Retrieval of Emotional and Non-Emotional Information. In B. Uttl, N. Ohta, & A. L. Siegenthaler (Eds.), Memory and Emotion Interdisciplinary Perspectives. Oxford, UK: Blackwell Publishing.

Uttl, B., Henry, M., Uttl, J., 2007. Human Factors in Avalanche Avoidance and Survival. Canadian Society for Brain Behavior and Cognitive Science, Victoria, BC.

Uttl, B., Uttl, J., & Henry, M., 2008. The Avaluator Avalanche Accident Prevention Card: Facts, Fictions, and Controversies. International Snow Science Workshop, Whistler, BC.