

## WORDS OF ESTIMATIVE PROBABILITY AND THE LANGUAGE OF THE NORTH AMERICAN PUBLIC AVALANCHE DANGER SCALE. ARE WE ALL COMMUNICATING THE SAME RISK?

Jimmy Tart<sup>1</sup>

<sup>1</sup>Park City Snow Safety, Park City Powder Cats, Ski Arpa, Park City, Utah, USA

**Abstract:** I have been wondering about something for years; the North American Avalanche Danger Scale uses words of estimative probability such as “certain” and “possible” in the avalanche forecast in regard to the likeliness of triggering an avalanche. Are we really in agreement on what these words mean? To help answer my question, I created a study that shows this variance in perceived values. My study consisted of five questions regarding the perceived probabilities of each of the following words from the Likelihood of Triggering section of the North American Public Avalanche Danger Scale; “certain”, “very likely”, “likely”, “possible” and “unlikely”. In addition to these five questions, there were three questions to establish the respondents’ profession and experience with statistics and probability. The distribution of the questionnaire was intended to reach not just Avalanche Professionals and recreational backcountry avalanche terrain users, but also segments of the professional world that involve risk management and that have embraced probabilistic thinking. Doctors, nurses, securities traders, pilots and firefighters were all targeted in the initial distribution.

Survey responses showed that while we humans all are thinking similarly about the probability attached to the word “certain”, as the scale descends toward “unlikely” the range of responses increases and the distribution of answers widens regardless of profession or level of training. Furthermore and most interesting, professional avalanche workers don’t have a more uniform distribution of responses than do recreationalists, medical professionals or people with formal training in statistics and probabilities.

I propose that using data analysis from this study and an adoption of practices common in the intelligence analysis world we could advance our collective communication skills.

More specifically we could improve our collective risk assessments and communication by fostering agreement on the words of estimative probability that are key to using the current North American Public Avalanche Danger Scale.

**KEYWORDS:** Forecast, Risk Communication, Behavioral Economics

This all started with a conversation with a co-worker. After completing our avalanche mitigation routes, it was routine for our supervisor to ask over the radio if we were done and ready to open the area. A common response among our mitigation teams was “I feel good about it” and then we would call our area clear to open. One morning after opening our area, my coworker and I got into a discussion about how vague that response actually was and how “feeling good” for one person might hide a different risk tolerance and level of uncertainty that another person might have.

The language that we were using, and possibly misusing, to try to describe our feelings employed single words or terms that are known as “Words Of Estimative Probability”.

Used to convey the likelihood of an events occurrence, some more common of examples of Words of Estimative Probability (WEP’s) appear in the “Likelihood of Avalanches” section of the North American Public Avalanche Danger Scale (NAPADS): “certain”, “very likely”, “possible” and “unlikely”. We all know what these words mean and we all know how to use them in a sentence. However, I had a suspicion that the numerical probabilities that we attach to them and our corresponding risk assessments were not uniform among professional avalanche workers or recreational winter backcountry users or even humans in general. So I designed a survey and to show if we all actually are in agreement on what these words mean. Having the respondents quan-

---

*Corresponding Author* : Jimmy Tart  
Park City Powder Cats, Ski Arpa  
(435) 901-4831  
thejimmytart@gmail.com

tifying the words of estimative probabilities from the danger scale can give us data to further refine our personal and public risk communication.

### 1. Survey Methodology

To establish if the respondents experience, the survey had three questions. The first asking the respondent to self-identify as either a professional avalanche worker, someone who recreates in avalanche terrain, or “none of the above”. The second question asked respondents if they had formal training in probability and statistics and/or used it in their work. The third question that was used to sort the responses was just “What is your profession?” with a space for a short answer.

The five remaining questions were all asking the respondents to place a single numerical probability on the WEP’s in the “Likelihood of Triggering” section of the NAPADS (certain, very likely, likely, possible, unlikely) with a full view of the North American Public Avalanche above the questions for reference (fig 2).

Groups target for distribution via social media included not just ski patrollers, guides and forecasters, but also pilots, firefighters, casino dealers, financial analysts and securities traders.

### 2. Results

There were 879 responses to the survey. One respondent who listed zero probability for all of the questions was eliminated from the results.

Figure 1 shows the distribution of responses to the survey questions about the words of estimate probability. First, the middle-quartile responses are distributed in an orderly, step-like manner. The WEP “certain”, which came first in the questions as noted had a nice, tight range of middle quartile responses with 77% of the respondents ascribing it a value of between 95% and 100% and 90% ascribing it a value of between 88% and 100%.

However, as we move to the other WEP’s the range of the middle-quartile responses grows gradually, implying less agreement about what the WEP’s imply in a probabilistic sense.

Outside if the middle quartiles is where things get interesting. What we see here is often a meaningful number of respondents who define a value one step, sometimes more, differently than the majority of respondents. For example: Virtually the entire range of middle quartile responses for “possible” lines up with the first quartile range of responses for “likely” and the fourth quartile of “unlikely” and the fourth quartile responses of “possible” match the range of the first three quar-

tiles of “very likely” (fig 1). In other words, while the responses in the middle quartiles actually line up quite nicely, the first and fourth quartile respondents have a range of responses that are a probabilistic Tower of Babel.

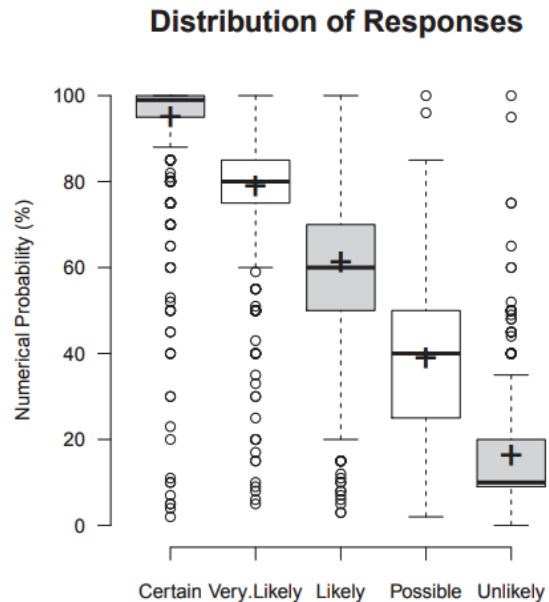


Figure 1: Distribution of Responses to Survey on Words of Estimative Probability

Center lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; crosses represent sample means. N = 878 points per sample. It should be noted that in this case the inclusive range between the 25th and 75th quartiles encompasses as much as 78% of the responses for each sample. This is due to the fact that the responses to the survey were largely skewed towards multiples of 5 (ie 70%, 75% etc)

### 3. Interpretation of Results

It is pretty clear that there is, at best, a reasonable amount of agreement at what these words mean. But at worst and on the margins, many of us are talking about completely different values. A disconcerting point made by that data is the large overlaps in perceived values between the WEP’s. At times the responses for each WEP fell into the overlap of one and sometimes two different WEP’s. Some of this drift could be attributed to the format of the questions in the survey. “Certain” came first and 100% is a good place to start to answer the questions in a descending, logical fashion. Regardless, there are accidents to be prevented in the data.

One hypothesis was that professional avalanche workers are in better agreement about how to define these words than the recreational users. The results of the survey do not support

this conclusion. The professional user group is not consistently in better agreement about the words of estimative probability than the recreational user group.<sup>1</sup>

All in all, the data seems to show that there is potential to improve the communication of risk through a better understanding of how we collectively define the words of estimative probability in the NAPADS. Because the use of these and other WEP's is common in avalanche risk communication amongst both professional and recreational avalanche risk communication, using this data to tighten the ranges and overlaps in the perceived values of the WEP's might be a good way to reduce avalanche accident numbers while also highlighting the limitations of WEP's.

#### 4. What to do with this?

Currently in North America, the accepted tool for the job of communicating the backcountry avalanche hazards and risks is the NAPADS and one of the elements in the NAPADS is an assessment about the likelihood of avalanches using words of estimative probability. Looking outside our industry for solutions to the question of how to communicate risk could help guide us to a refinement of the NAPADS. Further understanding about what is being said and what is being heard, combined with some understanding about the limitations inherent in words of estimative proba-

bility could allow people to refine their personal risk analysis and hopefully reduce the number of avalanche accidents.






#### 5. The Wisdom of Crowds

“The Wisdom of the Crowds” by James Surowiecki details, among other things, how aggregating the opinions of a large enough crowd of independent individuals can yield a reasonably accurate estimate value for an estimate.

So using this data for a “Wisdom of the Crowds” style analysis by which we could come to a midpoint in the perceived values of the WEPs could improve our collective communication skills and risk assessments by fostering agreement on what we are communicating with the current North American Public Avalanche Danger Scale and therefore possibly reducing the number of close calls and avalanche accidents in backcountry avalanche terrain.

#### 6. Words of Estimative Probability

“Words of Estimative Probability” by Sherman Kent of the CIA was originally published in 1964. In it, Kent eloquently lays out the challenges then faced by the CIA with risk communication which closely resembles the challenges faced by the avalanche community today. In his paper, he relates a story in which he and one of

North American Public Avalanche Danger Scale				
Avalanche danger is determined by the likelihood, size and distribution of avalanches.				
Danger Level		Travel Advice	Likelihood of Avalanches	Avalanche Size and Distribution
<b>5 Extreme</b>		Avoid all avalanche terrain.	Natural and human-triggered avalanches certain.	Large to very large avalanches in many areas.
<b>4 High</b>		Very dangerous avalanche conditions. Travel in avalanche terrain <i>not</i> recommended.	Natural avalanches likely; human-triggered avalanches very likely.	Large avalanches in many areas; or very large avalanches in specific areas.
<b>3 Considerable</b>		Dangerous avalanche conditions. Careful snowpack evaluation, cautious route-finding and conservative decision-making essential.	Natural avalanches possible; human-triggered avalanches likely.	Small avalanches in many areas; or large avalanches in specific areas; or very large avalanches in isolated areas.
<b>2 Moderate</b>		Heightened avalanche conditions on specific terrain features. Evaluate snow and terrain carefully; identify features of concern.	Natural avalanches unlikely; human-triggered avalanches possible.	Small avalanches in specific areas; or large avalanches in isolated areas.
<b>1 Low</b>		Generally safe avalanche conditions. Watch for unstable snow on isolated terrain features.	Natural and human-triggered avalanches unlikely.	Small avalanches in isolated areas or extreme terrain.

Safe backcountry travel requires training and experience. You control your own risk by choosing where, when and how you travel.

<sup>1</sup> F-Test and Brown-Forsythe tests for equality of variances each showed that the difference in variances between the professional group and the recreational group were significant ( $\alpha=.05$ ) for only one out of the five words of estimate probability.

his coworkers have a discussion regarding the numerical probability of a WEP that Kent used in a National Intelligence Estimate regarding the invasion of Yugoslavia by the Soviet Union in 1951. Kent sums up our problem nicely by writing that: “There is a language for odds; in fact there are two - the precise mathematical language of the actuary or race track bookie and a less precise though useful verbal equivalent. We did not use the numbers, however, and it appeared that we were misusing the words”.

Mr. Kent spends the rest of the paper detailing his efforts to codify the language of WEP's used by the intelligence community in the mid-1900's.. It seems his tactic of dictating approximated numerical probabilities and assigning these probabilities to an ever bloating list of WEP's was met with too much resistance from his colleagues. His efforts were not able to change the culture of communication in his community. The lesson is that trying to dictate definitions of WEPs is not effective. Noting that a heavy-handed bureaucratic decree seemed to fail indicated a more subtle path might be better.

### 7. Handling and Mishandling Words of Estimative Probability

A second example of the use to WEPs in CIA communication is a paper titled: “Handling and Mishandling Estimative Probability: Likelihood, Confidence, and the Search for Bin Laden” by Jeffrey Friedman and Richard Zeckhauser. In it, the authors outline the communication of probabilities of Osama Bin Laden's whereabouts by the US intelligence communities to then President Obama.

By 2015 the culture of risk communication seemed to have changed somewhat at the CIA. Their paper makes stating probabilities in numerical probabilities and not using vague WEP's sound more commonplace. However, Friedman and Zeckhauser do note a couple refinements that could be made in the communication. Notably, that the analysts would commonly state a range of possibilities (i.e. 20-40% chance) rather than a precise point and that confidence level or uncertainty in an analysis was something that had been implied by using a WEP, but was missing in a numerical probability.

Also, in arguing their case, they refer to the economist Daniel Ellsberg who wrote about what he called “ambiguity aversion” and wrote about in his paper “Risk, Ambiguity and the Savage Axioms” published in the *Quarterly Journal of Economics*. In it, Ellsberg designed a study that showed that people prefer to bet on odds that are established as opposed to ambiguous ones fur-

thering the case for having recognized values for the WEP's.

### 8. Nudge

Thaler and Sunstein wrote their book “Nudge” about their accumulated knowledge in the field of decision making from an Economics perspective. With decades of work studying the hows and whys of human decision making, Thaler and Sunstein finally were able to lay out some of the psychology behind steering humans to make better decisions. They define the role that we as risk communicators have as that of *Choice Architects*, a Choice Architect being someone who “has the responsibility for organizing the context through which people make decisions”. As Choice Architects, we have the ability, through the careful design of our communication, to “Nudge” the consumers of our risk communication to better decision making. To Thaler and Sunstein a Nudge is “any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere Nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting the fruit at eye level (in a cafeteria line) counts as a Nudge. Banning junk food does not.” Forecasters framing the forecast around what the study shows us that the majority of people think the WEP's mean as opposed to dictating to the majority what they should think would be a Nudge. As for the opinions outside of the majority, education of the problem might be the route to dragging them into the mainstream and narrowing the range of responses.

Nudging works. Both our federal government under President Obama and the government of the UK have set up so-called “Nudge” units in order to put the science of Behavioral Economics to practical use by increasing the performance of various departments. Obama formed the Social and Behavioral Sciences Team in 2015 and the UK has had their Behavioral Insights Team (BIT) since 2010. The BIT in the UK has shown particular success with their methods, with efforts ranging from improvements in the timeliness of tax payments to increases in voter registration. Their documented successes have shown these techniques and tactics can work.

In this context, the goal of the NAPADS is not to keep people away from avalanche terrain. The goal of any avalanche forecast is to help its consumers make better decisions about when to be where. Reducing the lack of agreement on the WEP's might not only reduce our aggregated exposure to the avalanche hazards on higher dan-

ger days, but allow for less uncertainty about the risks in avalanche terrain on lower danger days and consequentially greater safety in grander objectives. So to Nudge forecast consumers is not to keep them out of certain areas, but to give them better tools to decide when to be where and the greater freedom to travel in the backcountry.

## 9. Summary

The whole goal of this project is to Nudge Decision Makers in the winter backcountry to make better decisions by framing the Words of Estimative Probability used in the North American Public Avalanche Danger Scale in terms of numerical probabilities produced by an analysis of the survey data. Not as an effort to dictate to the public what they should think, but by telling them what they said they thought. Although the numerical probability for the likelihood of triggering an avalanche on any given slope would be realistically impossible to ever state to any degree of accuracy, simply understanding our perceived values of the Words of Estimative Probability would allow the North American Public Avalanche Danger Scale to more clearly give Decision Makers the tools with which to make a better decision regarding terrain and timing.

## Acknowledgement

Big thanks to Mark Saurer for setting the standard, Paul Diegel for being persnickety on the language of the survey, Allie Fredbo for knowing what an abstract is and Doug Krause for his infectious curiosity.

## References

Statistical analysis courtesy of Ken Thompson with Chione Analytics.

Figure 1:

Data: Tart, James (2018). Survey of Words of Estimated Probability as used in the North American Public Avalanche Danger Scale. Unpublished raw data. Graphics: Tyers and Rappsilber labs (2014). Boxplot generator using R [web-based software]. <http://shiny.chemgrid.org/boxplotr/>

Ellsberg, Daniel Nov 1961: "Risk, Ambiguity and the Savage Axioms", Quarterly Journal of Economics

Freidman Jeffrey A. & Zeckhauser, Richard 2015: Handling and Mishandling Estimative Probability: Likelihood, Confidence and the Search for Bin Laden, Intelligence and National Security, 30:1, 77-99

Kent, Sherman, 1964: "Words of Estimative Probability" Sherman Kent P49-64 Studies in Intelligence

Surowicki, James, 2004: The Wisdom of Crowds, Doubleday, Anchor 336pp

Thaler, Richard H. and Sunstein, Cass R., 2008: Nudge: improving decisions about health, wealth and happiness, Yale University Press 312pp