[Methodology Overview 1](#_Toc39408524)

[Study Period 2](#_Toc39408525)

[Notes on AAC report narratives, data quality, and sampling issues 2](#_Toc39408526)

[Data cleaning steps 3](#_Toc39408527)

[Types of Factors 4](#_Toc39408528)

[Selected Dictionaries of Terms and Category Notes 4](#_Toc39408529)

[Omitted, Added, and Combined Categories Versus Traditional AAC Tagging System 4](#_Toc39408530)

[Additional Category Notes 7](#_Toc39408531)

### Methodology Overview

1. Data sourcing: To analyze the American Alpine Club’s accident narratives, I copied the accident titles and narratives using the AAC website’s accident report search tool, choosing the publication year and then pulling the data from linked article pages displayed in the results. The project began in late 2018, and by the time the 2019 accidents were published, an AAC website project was underway that disabled the search by year function. Because of that, the AAC provided the PDF version of the 2019 accidents for analysis. One more note for anyone trying to find the accident reports is that as of April 2020, the AAC site has issues sorting by year.
2. Differences from the AAC methodology: As stated in the *Rock & Ice* article, this analysis makes significant changes to the traditional AAC categories listed in their tables: For example, accidents are tagged to disciplines like Top-rope, Sport, Trad, and Ice Climbing, data not tabulated by the AAC. Another example is that unlike the AAC, I didn’t attempt to identify one immediate cause plus contributory causes: All tags are handled as factors that contributed to the accident outcome. Finding associations between factors was the goal.
3. Categories tagged: I aimed to tag each report with accident factors, along with climber ages, experience levels, outcome severities, and month of the year. Many accident factors come from the AAC’s tables of causes.   
     
   A key difference from the AAC is that I didn’t attempt to identify a single immediate cause for an accident, and to separately identify contributory causes. Some categories the AAC uses were not used for this analysis (terrain), while other categories used in this analysis are not tracked by the AAC (climbing discipline). Much more on that in the category section below. My goal was to tag factors that would be actionable for climbers studying the data. In making category decisions, I consulted the AAC editors and climber and data analyst Nicholas Cohn-Martin.
4. Categories Tagging Methodology: Broadly speaking, I used Excel formulas to search accident titles and reports for lists of key phrases: “Dictionaries.” Generally, shorter lists of keyterms were needed for searching accident titles, which tend to be very defined by formulaic AAC styles. Longer dictionaries were used for accident descriptions, and were often revised to be tighter or looser as I worked. I tested the tags by looking for agreement between Title & Description dictionaries, which meant the strongest tag. I also considered if there was a contradictory signal (or no signal) from the other field or from a conflicting category. Additional review included arbitrating between tags for conflicting categories (initial run of Title & Description comparison, then use of Description scoring of how many terms turned up).

For every category, I reviewed the results to confirm the key terms were working. When they generated incorrect results, I dropped overly broad terms from the formula, or manually deleted incorrect tags. This iterative text analysis allowed me to revise my tags for categories, return to them when I found problems later, and improve my methods through the project.

1. Analysis method: Microsoft Excel’s automatic table create system, Pivot Tables, allows a user to highlight a range of data and sort and filter the results as desired without risking any changes to the source data. That allowed me to manipulate the data in various ways, for example, filtering out incidents that were labeled as “Non-climbing.” That filter, excluding “Non-climbing” incidents, is applied to all analyses in the *Rock & Ice* article unless otherwise noted.

### Study Period

The study period of 30 years starts with 1990 publication year reports (mostly 1989 incidents) and ends after with the 2019 publication year reports (mostly 2018 incidents).

That time range of 1990 to 2019 captures accident reports during the second wave of modern techniques and equipment (the first wave of modern climbing equipment, in decades leading up to 1990, includes introduction of specialized harnesses, shoes, and dynamic rope).

● 1977 - Cams hit market, are improved upon

● 1987 – Black Diamond Camalots sold as first double-axel cams

● 1990s - Bolting increases, sport climbing grows

● 1990s and 2000s - Climbing gyms take off

### Notes on AAC report narratives, data quality, and sampling issues

AAC reports vary significantly in level of detail, from a couple short paragraphs to many hundreds of words. Report origins vary from first-hand reports of climbers to park ranger reports to media reports. That means no analysis of these reports can be a perfect reflection of reality.

At the same time, the AAC reports represent a rich, detailed dataset of accidents in which the aim of the reporters and editors was to illustrate accident causes and caution other climbers. It is probably one of the best available datasets to attempt a natural language analysis like this.

Another key issue with AAC reports to keep in mind is that it is not a truly random sample:

First, not all incidents are reported, with less serious injuries even less likely to generate reports. Second, the AAC *Accidents* editors choose a selection of reports, not publishing them all in a given year, so only a subset is available for this analysis. Compared with the table counts, this analysis seems to have 56% of the total accidents reported to the AAC for the study period.

For much of the past 30 years, the AAC editors did aim to publish as many reports as possible. Over time as more reports have been sent in by volunteers they’ve become more selective, with a marked shift when Dougald MacDonald, executive editor for the American Alpine Club took over as editor in 2015, he says: “In the past, we tried to document every technical climbing and mountaineering accident, regardless of whether we could offer meaningful analysis. We are more selective now with a focus on the educational value of a narrative.”

Lastly, outdoor bouldering and indoor gym accidents are rarely reported to the AAC, and so are underrepresented in the data. As explained by MacDonald: “This is largely because most injuries in gyms and bouldering are relatively minor … I believe that a very large majority of people who have bouldering injuries self-evacuate and get themselves to a doctor, if necessary, and thus there is no search and rescue report to submit.”

### Data cleaning steps

Duplicates cropped up when processing the AAC online entries. Besides exact duplicates, I also had to clean aggregate reports. Aggregate reports fall into one or both of two basic categories.

The first type is in-depth incident descriptions with a note appended about other similar accidents at the same location. Generally, those additional descriptions need to be ignored for tagging purposes. I’ve attempted to isolate all of those (marked “none, truncated”) and delete extra details, but will restore the complete narratives in the final step of assembling a cleaned and tagged database. Short addendums that only describe similar accidents with the same characteristics can be left intact.

The second type is shared reports—reports on one or more incidents, as from the Gunks or Denali. I’m pasting those all into a separate sheet, with the unique identifiers preserved. Where there is an incident or are multiple incidents with decent detail within those reports, I will keep them in the sheet, adding A, B, C to the incident ID. The standard for keeping split incidents is to have at least three factors mentioned; or a good climber description plus factors; or a long explanation; or in the case of a fatality with a clear cause, keep that (especially a soloing death where cause may never be known anyway in a longer report). A T added to the end of the ID means the incident was truncated for categorization purposes.

Through several stages of category tagging, manually discovered duplicates that evaded prior notice were catalogued and cleaned up with this methodology.

*Incomplete Data* – Most incidents lack descriptions to assign a category for some factors. Sometimes there is partial information. In those cases I aimed to leave the category untagged, aiming for greater accuracy and less completeness rather than the reverse. It’s an area of data cleaning that can continue to improve with further research—routes could certainly be tagged as single or multi-pitch based on guidebook or Mountain Project descriptions (some Mountain Project lookups were performed to categorize sport and trad and alpine routes, but not all). Accident dates could be found with historical searches. Missing experience levels could be filled in from people with knowledge of the climbers involved. I didn't attempt that level of research with this initial attempt.

### Types of Factors

The major types of categories I tagged are: Type of Climbing, Severity of Outcomes, Experience Level, Age, Human Factors, Natural Factors, Alpine & Ice Climbing Factors, Alpine & Ice Climbing Natural Factors, and Months.

Some tags are considered activities. These include: Lead climbing, plus Belaying / Rappelling / Lowering. That means I attempted to tag these as a “time-of-accident” status rather than applying the tag to all narratives that included verbiage that described the activity in a mundane way earlier in the hours leading up to an accident.

### Selected Dictionaries of Terms and Category Notes

For full, final formulas and notes about categories, see the **Category Dictionaries+ Notes** Excel file. For selected categories of note, read on:

### Omitted, Added, and Combined Categories Versus Traditional AAC Tagging System

**Terrain - Omitted**

Rock, Snow, Ice, River, and Unknown are the traditional descriptors for terrain. For this project the main goal was to explore accidents as defined by type of climbing, and to identify factors to watch out for with a high rate of association with other factors.

**Climbing Disciplines - Added**

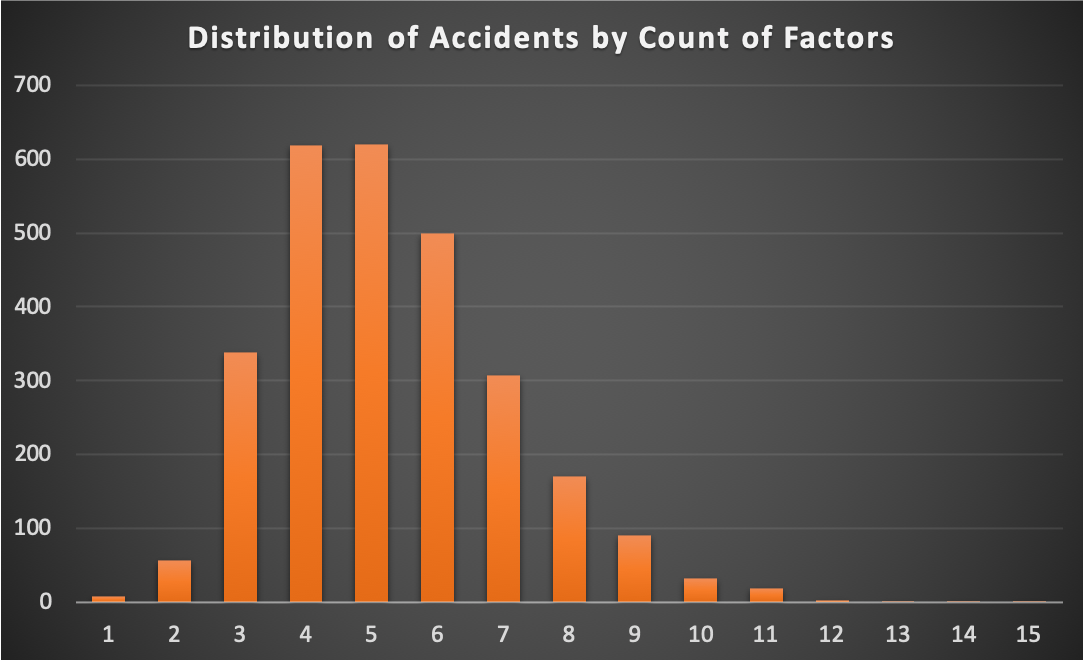
Instead of terrain types, I set the following “Types of Climbing” to categorize incidents: Single Pitch, Multi Pitch, Top-rope, Sport / Lead, Trad, Bouldering, Gym, Alpine / Mountaineering, Ice Climbing, Soloing / Alone. Some of these options are not exclusive, as noted in the Categories Excel sheet tab.

**Fall as an accident cause – Omitted**

For the same reason that I didn’t attempt to tag terrain, I didn’t try to tag falls. Falls represent the vast majority of Immediate Causes in the AAC tables: 6,954 that I see in the tables out of 9,083, as of the 2019 *Accidents* report. I didn’t count falls because I see them as a normal part of climbing. Tagging falls would put them at the top of the list of causes no matter how you slice the data, which I believe would obscure more actionable factors to watch out for. And, those other factors are what take a fall from being a normal part of a climbing day to becoming an “accident.” If someone falls but there is no belay error, it might not be an accident. The intent was trying to tease out the factors responsible for why the accident was an accident, not a “normal fall.”

Of course, in some cases normal falls can cause injuries and an accident, but most accidents in the database did receive multiple tags aside from the baseline categories of age, experience level, month of the year, and outcome severity, averaging 4.3 factor tags per incident.

The distribution of 2,700+ accidents by count of tags received can be seen below.



**Immediate Cause / Contributory Cause Statuses - Omitted**

While many of these causes are the same as the AAC’s in this study, no attempt is made to assign the causes as immediate or contributory. The reason is that as I tag, my judgment may differ from the AAC editors in many cases, and even experts or people with direct knowledge of the incident will disagree. On top of that it seems to me that causation or primary importance of a factor is extremely subjective.

My goal is a behavioral database--not literally what surface someone fell on, but what were they doing? The intent is to create a database that is association-based, without attributing immediate and contributory causes. Causes are often interrelated and subjective, with contributory causes sequentially first in some cases. Would a fall be a primary cause, even if immediate, if Off route or Inexperience is involved? Since people fall all the time, the contributory causes of belay and protection seem more important than the immediate cause of a fall.

Instead, accidents are divided into Human Factors, Natural Factors, and Other Factors—again, containing many of the same categories.

**Sex / Gender - Omitted**

Sex is an AAC category factor for accidents, but it has rarely been applied, with only several hundred items in the 1951-present Accident Tables. So, already it is fairly unused. For my study I’m interested in examining likely causes of accidents such as conditions, experience, style of climbing, and specific mistakes to be avoided. Sex or gender aren’t of interest to me any more than weight or race are.

**Ascent - Omitted**

Far more incidents occurred during Ascent than Descent, so it is the normative case. There are many reports where Ascent is assumed but not confirmed versus traversing or resting. Skipped this category, but did tag Descents since it is more often called out and described as an element.

**Injury/Fatality Counts and Categories of Injury or Illness - Modified**

My methodology doesn’t lend itself to exact counts of injured parties, and since nearly all incidents either were serious threats to life or could have been, I’m skipping an attempt at recounting injured individuals one by one via narratives, which don’t always provide much detail.

Likewise the traditional injury types, fractures, contusions, etc. don’t illustrate severity. Frequently people have a grab bag of injuries, and it doesn’t define the incident in the same way as accident factor categories, but can just illustrate the venue/climbing type, once again. Injuries causing an accident (twisted knee ends a trip, shoulder displacement, ascent illnesses like HAPE and HACE) are categorized as they represent an entire accident/incident cause.

Instead of counts or injury categories, I’ve tried to classify the severity of each incident fatality (at least one), serious injury (one or more), or no injury or minor injuries only. The first two classifications aren’t exclusive, and the last one is.

**Ascent Illnesses - Modified**

Combined HAPE, HACE, AMS, “Ascending too fast,” into one category tag, because all are related altitude illnesses that climbers can take seriously as a group of undesirable issues that can be avoided by ascending at a conservative pace, and mitigated by listening to your body and communicating problems to partners. Ascent illnesses aren’t yet systematically tagged to Brain injury, although cerebral edema tagged as Ascent Illness may rightfully makes sense with both tags—researchers who want to take this approach can do so by combining categories or selectively adding incidents tagged to Ascent Illness.

### Additional Category Notes

**Experience**

If a party had a very experienced person and someone with little to no experience, and both were involved in a failed belay, both tags were assigned. In cases of a singular person causing an accident with others present, their experience levels were not tagged. This optimizes the data for accident causality and presents a shortcoming if someone were to study the data looking for success of post-accident rescue attempts, with nearby party experience relevant.

Tags are meant to describe a person's experience at the climbing discipline for which the accident occurred. Someone with lots of experience but only a day of trad climbing, who gets in a trad climbing accident, would be considered to have No or Little Experience. In cases where the AAC account describes the person as experienced and learning a narrow skill or activity that is not a discipline of climbing, it isn't used to demote them to status of a beginner.

One change to the AAC system: “Moderate” may indicate 1-3 years of experience if the fact of the number of years is all that is mentioned. But Moderate is not strictly defined as 1-3 years of experience. Instead, No/Little, Moderate, Experienced, and Unknown are treated as qualitative rather than strictly up to 1 year, 1-3, and 3+ as that is a very rigid definition and doesn't cover mixed experience. Two years of sport climbing for one person with 3x a year might be little experience, and 6 months of sport for another who dirtbagged around climbing 5x a week could be "Experienced." Qualitative descriptions are thus used when in the reports.

**Age**

Because AAC reports include many ages for many parties surrounding an incident, it would be extremely difficult to manually go through all the entries to identify only the ages directly involved in the incident. I haven’t done that in my study, and so my initial age data is probably less reliable for attributing accident causality than Experience, for which I am reviewing entries and aiming to restrict Experience tags to people directly involved in an incident, not rescuers.

**Months**

In cases of multiple months mentioned in the accident report, the month tagged is the one in which an incident—accident or injury—took place, or when unclear, the more likely of two months mentioned when it probably did.

**Accident & Injury Severity**

The three outcomes for this analysis are Fatal, Serious Injury, and Minor or No Injury. Category overlap is allowed where appropriate, when multiple parties were injured to varying degrees.

The dividing line between Serious and Minor or No Injury can be subjective. For this analysis, loss of consciousness, broken bones, rescue needed, and hospitalization were all considered sufficient to be Serious. Minor and No Injury usually meant lucky escapes, and incidents where bruising and scrapes didn't result in any major rescue or medical problem. Some incidents explicitly describe loss of consciousness or hospitalization as "minor injury"--unless the hospitalization confirmed mere bruising and scraping, incident descriptions are over-ridden and broken bones and loss of consciousness are still considered serious. Introducing more delineation beyond these three states was considered but rejected as it would likely increase the need for subjective judgment calls.

During cleaning, tried to delete extraneous Serious Injury tags for victims who later died, yielding to the Fatal tag. Not deleting the Serious Injury tag when there are multiple party members and one is seriously injured, regardless of whether another party member dies. But more broadly, may be redundant serious injury tags in fatal cases. Could accept the inflated tagging numbers, or go back and review non-Hypothetical-tagged Fatals (not duplicating earlier review work).

When parties are described as uninjured to delineate them from the injured climber, and not because they were involved in a fall or active part of the incident that threatened health, the Minor and No Injury category was untagged.

During data cleaning, attempted to delete extraneous Serious Injury tags for victims who later died, yielding to the Fatal tag. Many incidents do legitimately contain tags for multiple levels of injuries because of varying outcomes of different victims.

**Alpine / Mountaineering Notes**

Note related to Ice Climbing: Some incidents get both the ice climbing and mountaineering tag. Ice Ax is not by definition considered ice climbing, because it is of use in mixed alpine environments for snow and slick surfaces alone.

**Roped and Unroped**

Related to Roped/Unroped categories: Mountaineering and alpine climbing often involves a mix of roped and unroped climbing, with some free climbing and some rappelling. For incidents where it’s unclear that either was a key factor, or was the situation at the time of the accident, neither the Roped nor the Unroped tag is applied. Generally, Roped or Unroped is assessed for the activity going on at the time of the accident.

**Trad Climbing**

Setting protection to rappel, like a nut, chockstone, cam, etc. is counted as involving trad climbing when those tags were reviewed. When aid climbing was tagged as trad, I left it that way on review. Ice Climbing with ice screws and gear only is not considered trad, so that different risks can be analyzed. In cases of mixed routes that include ice as well as nuts placed in cracks, the aim would be for both tags to be applied.

**Rushed**

Borderline cases: Rushed is tagged in cases where haste is mentioned as likely; when listed as a hypothetical cause with other alternative unconfirmed factors, it is not tagged.