Since its introduction in 2006, the Extended Column Test (ECT) has become one of the most popular tests to assess point snow instability. In 2009, two studies explored how ECT results correlated to observed instabilities, laying the foundation for the tests’ interpretation we are using today. Based on data from the first winters, Ron Simenhois and Karl Birkeland showed that ECTVs and ECTXs were typically observed when conditions indicated instability, while ECTNs and ECTXs were mostly observed on stable slopes (for ECT scoring abbreviations refer to observational guidelines [Greene et al., 2010]). This continues to be the standard for interpreting ECTs in the United States. In Switzerland, Kurt Winkler and Jürg Schweizer noted that ECTPs21 detected a large proportion of unstable slopes correctly while keeping the number of false alarms low. Again, ECTNs or ECTXs were more frequently associated with stable slopes in their study. In Switzerland, this is the operationally used approach to classify ECT results.

Now, more than ten years later, the ECT is a well-established test internationally. The time is right to revisit these stability interpretations, as recently done using Swiss data (Techel et al., 2020), by combining ECTs from North America (mostly from the U.S.), Spain, and Switzerland.

WHAT DATA DID WE USE?
We explored several snow profile databases from snowpilot.org, Val d’Aran (Spain) and Switzerland. We only included backcountry snowpit profiles with ECT results and information about the presence or absence of clear signs of instability. In total, we had:

• 2,579 ECTs from snowpilot.org, with about 90% from U.S. (snowpilot.org is open to the public)
• 167 ECTs from Val d’Aran / Spain, with profiles mostly collected by forecasters and observers
• 1,226 ECTs from Switzerland, with profiles observed by researchers and field observers

These ECTs are therefore just a small subset of the more than 30,000 combined ECTs in these databases.

HOW DID WE ANALYZE THE DATA?
For each ECT, if more than one failure was indicated we used the following rules to decide which result was the most relevant for stability assessment.

1. If an ECTV or ECTP failure was recorded: we considered the lowest number of taps required for full propagation.
2. If full propagation was not observed, we considered the lowest number of taps associated with the ECTN or ECTX.

If there were several ECT results in the same snow pit, we randomly picked one. This provided us with a dataset of almost 4,000 ECT results.

WHAT DID WE FIND?
Quite clearly, ECTVs and ECTPs are observed more often on unstable slopes (red line in Figure 1 located above the base rate, represented by the dashed black line), while ECTNs and ECTXs are observed more commonly on stable slopes (yellow line located below the base rate) (Figure 1). Further, ECTs with a higher number of taps tend to be more stable. ECTPs with less than 14 taps were the most unstable, with about 60% of those tests being associated with avalanches or signs of instability. This is about double the number of locations associated with avalanches or signs of instability in our entire dataset (the base rate). While still clearly on the unstable side of the base rate, the proportion of unstable locations decreases with more taps, even with an ECTP result. When more than 22 taps are necessary to initiate a fracture in an ECTP the proportion of unstable slopes was not significantly higher than the base rate, indicating that such results might be linked to something like “intermediate” stability. We note a similar result for ECTN>8, while ECTN>8 was clearly linked to stability.

INTERPRETING THE FINDINGS
In a perfect world, we would know absolutely whether a slope can be triggered or not. However, in reality, all studies exploring stability tests—including this one—must use other observations to infer slope stability. If the slope stability rating is wrong, which is inevitable for at least part of our data, then the test accuracy drops. For example, in our study we likely had at least some cases where observers did not see any signs of instability but the snowpack was still unstable and avalanches could be triggered. Similarly, there are also likely cases where observers noted signs of instability on nearby slopes, but the slope being tested was in fact stable. These situations lead to a misclassification of the slope stability and have the potential to lower the correct classification by the stability test being evaluated. However, while these cases influence absolute values, it does not influence the observed patterns in Figure 1. We can see this when we compare our much smaller Spanish data set, which was thoroughly quality-checked by the forecasters in Val d’Aran, to our U.S. and Swiss data sets, which both relied on observations submitted together with snow profiles. In Spain, the proportion of unstable locations was about 80% for ECTP>23, and 8% for ECTN and ECTX in a data set with 35% unstable slopes (Figure 2b). In the U.S. and Switzerland, absolute values and the shape of the curves were remarkably similar (Figures 2a and 2c). The only difference was that the proportion of unstable slopes for ECTP>22 was slightly above the base rate in the U.S. and slightly below in Switzerland.

TAKE-HOME POINTS
The correlation between signs of instability and ECT scores clearly shows that the ECT is a valuable test for assessing snow instability. Our data confirms the findings in the Swiss study that...
The correlation between signs of instability and ECT scores clearly shows that the ECT is a valuable test for assessing snow instability.

Figure 2: Proportion of unstable ECT locations for each combination of fracture propagation and number of taps until failure for the three data sets. The Snowpilot (a) and Swiss (c) results, which are based on a large number of ECT, look rather similar. In contrast, the ECT data from Spain discriminates better between ECT results indicating instability and failure for the three data sets. The Snowpilot (a) and Swiss (c) results, which are based on a large number of ECT, look more similar. In contrast, the ECT data from Spain discriminates better between ECT results indicating instability and failure for the three data sets.

Figure 3: Relating ECT results to observed signs of instability in the surroundings in this data set. The stability class poor is split into two sub-classes, reflecting the trend seen in Figure 1 for an intermediate number of taps.