



MS. TECH | EDISON: LIBRARY OF CONGRESS, BULB: PIXABAY

Business Impact**How the data mining of failure could teach us the secrets of success**

These data researchers found that for startups, scientists, and terrorists alike, learning too little from experience spells doom.

by Emerging Technology from the arXiv March 29, 2019

Thomas Edison is often described as America's greatest inventor. His successes include electric power generation, sound recording, and the electric lightbulb.



But Edison was no stranger to failure. He famously tested 1,000 different designs before settling on the carbon filament that became the first commercially successful lightbulb. This tenacity set him apart. “Many of life’s failures are people who did not realize how close they were to success when they gave up,” he said.

Many groups and individuals have studied the nature of success. These studies have yielded varying degrees of insight. The flip side—the nature of failure—is much less well studied but arguably more important. Little is known about the mechanisms that govern the dynamics of failure.

Today that changes, at least in part, thanks to the work of Yian Yin at Northwestern University in Evanston, Illinois, and colleagues. This team has analyzed the nature of failure in three huge data sets following the fortunes of startup companies, researchers attempting to secure funding, and terrorist attacks. The work reveals the dynamics of failure and a hidden signature that can separate impending failures from successes at an early stage.

The team’s method is based on the analysis of three data sets. The first is a set of all health-related research proposals submitted to the US National Institutes of Health between 1985 and 2015.

The NIH is the world’s largest funder of biomedical research, so this data set is huge, consisting of 776,721 applications by 139,091 researchers. It also includes information about whether or not each proposal was funded; in other words, whether or not it was successful.

The second database is of investment records in startup companies from VentureXpert, the official database for the National Venture Capital Association. This follows the fate of every startup funded by venture capitalists between 1970 and 2017—a total of 58,111 companies involving 253,579 innovators.

In this case, a startup is considered successful if it achieved an initial public offering or high-value merger and acquisition within five years of its founding.

The final data set is from the Global Terrorism Database, which records 170,350 terrorist attacks by 3,178 terrorist organizations between 1970 and 2017. In this case, a successful attack is one that claims at least one life, while failures are those that kill no one.

A key feature of these data sets is that they allow Yin and co to follow the fortune of researchers, innovators, and terrorist groups that make numerous attempts to achieve their goal. A key question that they investigate is how attempts change over time and what factors are involved in these changes.

Yin and co specifically study two factors that are thought to play an important role in success and failure: chance and learning. They first look at chance, the notion that random events play an important role to hinder or boost the chances of success.

That leads to a simple model. If chance is the key factor that determines success, then each attempt has a finite probability of being successful. Indeed, success will eventually occur if enough attempts are made. This suggests that the number of attempts before a success should follow an exponential distribution.

To test this theory, Yin and co studied the sequences of failures by the same individuals or teams before they achieved a success. It turns out that these sequences do not follow the kind of distribution predicted by a chance model.

Yin and co also evaluated the first and penultimate attempts in these failure streaks and then compared them to see how they have changed. If luck is all that matters, there should be no significant difference.

But the penultimate efforts are significantly better than the first attempts, say the team. This suggests that another mechanism must be at play: the people involved must be learning. In other words, the experience of failure teaches valuable lessons that can be used to improve performance the next time around.

Since learning should reduce the number of attempts required before achieving success, it should lead to a narrower distribution of failure streaks than the exponential form predicted by the chance model.

But to the surprise of Yin and co, failure streaks do not follow this pattern either. In fact, they have a much fatter-tailed distribution. “These observations demonstrate that neither chance nor learning alone can explain the empirical patterns underlying failures,” the researchers say.

So what other factors are important? To find out, Yin and co modeled the way people learn from experience and how this influences their next attempt. In particular, they modeled whether people take into account all their previous experiences or just some of them.

The resulting model considers a complete range of learning—from agents who take all their past experience into account to those who do not take any of their past experience into account, and everything in between.

The team say the model predicts a phase change in the behavior that matches the empirical data. When the level of learning from experience is below some threshold, future attempts never become good enough to succeed. Indeed, groups can end up reducing the quality of their work.

But when the level of learning from experience is above this threshold, future attempts become better and better until they eventually succeed. And the key factor is the way people learn.

That has important implications. For example, it means that a team’s learning process is a good indicator of whether or not it will succeed at some point. “Our findings unveil identifiable yet previously unknown early signals that allow us to identify failure dynamics that will lead to ultimate victory or defeat,” say Yin and co.

The next step will be to analyze successful learning *in situ* so that it can be distinguished from unsuccessful learning and eventually taught systematically.

That could be a crucial way for teams to get an edge on the competition. And with so much at stake in terms of funding and investment, successful learners have plenty of incentive to try harder. Edison would surely be impressed.

Ref: arxiv.org/abs/1903.07562: Quantifying Dynamics of Failure Across Science, Startups, and Security