Last Words

Failure Is an Orphan (Let's Adopt)

Stan Szpakowicz*
SITE, University of Ottawa &
ICS, Polish Academy of Sciences

Sweet Smell of Success

We like to think that the progress of science follows an exponential curve, boldly shooting upwards, ever more speedily. In reality, there is a firmament of darkness (that's our daily slog) sparsely spotted with dim lights (those are discoveries). Fittingly, discoverers are also few, just as in every field of human endeavor. For every great leader, artist, or scientist there is a multitude of followers, and this is as it should be. There can only be so many winners—and only so much winning research. Anecdotal evidence has it that one project in a hundred gives back: Its results justify the investment.[1]

And yet, our cocky attitude: We tend to count ourselves among that one hundredth. We need success. We simply must excel. That is, after all, how scientists are judged. Those of us who have sat on hiring or promotion committees know how evaluation works; surely, we all get to be at the receiving end. One needs many publications for a thriving career, and we publish if we show how we go one better on someone else's result. So, we cannot afford a setback, a path wrongly taken, a poor result. There is no room for that in our line of work.

This already sounds dismal, but it gets worse. People have come to equate every negative result with failure. Suppose you have set up an experiment carefully and in good faith, but still it comes up short. That's not a positive outcome. Maybe your intuition has let you down. Maybe this *cannot* work. Wait, maybe you can *prove* that it cannot work? No, forget it. You have already failed. Don't waste any more time. Cut your losses and move on.

Well, reconsider! You can miss a truly fundamental lesson. In 1887, Michelson and Morley set out to prove the existence of ether, thought to be a conduit for light. They could not prove it: they failed. But—their null result prompted a line of research which culminated in the theory of special relativity.

House of Peers

Peer review breeds mediocrity. Galileos, Brunos, Modiglianis, and van Goghs go against the grain; the scorn of their peers and the indifference of the public crush them. An obliging self-promoter will fare much better. Peer review also censures failure. A

---

* School of Information Technology and Engineering, University of Ottawa, 800 King Edward Avenue, Ottawa, Ontario, K1N 6N5, Canada. *Institute of Computer Science, Polish Academy of Sciences, Ordona 21, 01-237 Warszawa, Poland.* E-mail: szpak@site.uottawa.ca.

1 As Einstein famously said: "I think and think for months and years. Ninety-nine times, the conclusion is false. The hundredth time I am right."

forthright admission of the inferiority of one's results—despite the integrity or novelty of the work—is a kiss of death: no publication. There must be improvement, however minuscule.[2] By the "law of averages," caution tends to trump adventurousness, and certainly conformance to reviewers' expectations is an asset. Indeed, we write so *they* are likely to accept.

Now, reviewers are us. If we the authors believe that negative results are always and inescapably a bad thing, so do we the reviewers. If we the authors relentlessly strive for success, we the reviewers check papers for signs of success. Serves us all right: If we have lost, we can't win.

Can't we?

Philosopher's Stone

Good science engenders good results. An experiment carefully thought out, a systematic procedure, an honest evaluation—these are the ingredients. It is not mandatory for the results to be positive, though it certainly lifts one up if they are.

In areas where empirical methods dominate—computational linguistics is such a discipline—it is a given that people try things which fail at the experimental stage. This may be due to lack of rigor, but often there are deeper, unexpected, and intriguing reasons. We can learn a lot if we analyze scientifically why an intuitive and plausible experiment did not work. We know how important counter-examples are in rejecting conjectures. And if a negative result points to an interesting and important open problem, it ought to be explored. Then again, to know what leads to dead ends in research surely can warn others off paths which take us nowhere.

An interesting negative result may arise from a compelling idea proven wrong in a thorough and well-conducted experiment. This is valuable to the community if you can explain confidently why you have tried out this idea, and if people can replicate your experiment; the setback is an educational experience insofar as you convince others that the expected outcome just could not materialize. Simply put, a negative result *can* be a useful lesson.[3]

---

2 A score will double if just a hundred researchers increase it, one by one, by a *mere* 0.7%.

3 Not everyone is loath to accept negative results. *Journal of Negative Results in BioMedicine* (`www.jnrbm.com`) welcomes articles which contradict current tenets, to better support scientists and physicians in making clinical decisions. *Journal of Negative Results – Ecology and Evolutionary Biology* (`www.jnr-eeb.org`) offers an outlet to studies which seem uneventful (as a reaction to publications in the field which only accept studies with noteworthy results) to avoid missing important phenomena in nature. *Journal of Articles in Support of the Null Hypothesis* (`www.jasnh.com`) publishes papers on experiments which do not pass the traditional significance level threshold, to avoid replicating empirical work already examined. Closer to home, there is WWWW: the AAAI Spring Symposia hosted in 2006 and 2008 workshops dedicated to "What Went Wrong and Why: Lessons from AI Research and Applications" (`www.aaai.org/Press/Reports/Symposia/Spring/ss-06-08.php`), `www.aaai.org/Press/Reports/Workshops/ws-08-14.php`. And if you want to publish a serious, worthwhile negative result in computational linguistics, there is *Journal of Interesting Negative Results in NLP and ML* (`www.jinr.org`).