Reflections on Lakatos's "Proofs and Refutations"[1]

Andrew Gelman

15 Oct 2020

My advanced math classes in college followed a standard pattern: in the beginning of the semester were the definitions, then came the lemmas, then the theorems, culminating at the end of the semester with the big proofs and then, if there was time, maybe some applications along with the much-despised "heuristics." And not a counterexample to be found. These were theorems, after all. A theorem is true and so has no counterexamples, right?

It was only in my senior year that I learned the proper order of mathematical reasoning: first the problem, then the theorem, then the proof, with the definition at the very end. The definitions come at the end because they represent the minimal conditions under which the theorem is true. The statement of the theorem itself changes as the proof and definitions develop. And, just as a country is defined by its borders, a theorem is bounded by its counterexamples, which are duals to its definitions.

I gained this perspective by reading Proofs and Refutations, a book taken from the Ph.D. thesis of the great philosopher Imre Lakatos. Lakatos's later work was in the philosophy of science, where his synthesis of the ideas of Karl Popper and Thomas Kuhn led to a sophisticated falsificationist model of scientific practice which influenced generations of social scientists. Proofs and Refutations finds a similar empirical sprit within mathematics.

The physicist Eugene Wigner wrote about "the unreasonable effectiveness of mathematics in the natural sciences." Flipping this around, Proofs and Refutations discusses the effectiveness of scientific inquiry in mathematical research, an idea which is commonplace in our modern era of computer experimentation but was controversial in 1964, when Proofs and Refutations was published, or even twenty years after that, when I was taking all those conventionally-constructed math classes.

In his book, Lakatos develops his ideas through a development of Euler's formula of the faces, edges, and vertices of polyhedra. It turns out this theorem has lots of counterexamples. One at a time Lakatos brings these up, in the setting of a fanciful conversation among hypothetical mathematics students, following the ideas of mathematicians working on this problem in the 1800s, and these hypothetical students manage to rescue the theorem from each counterexample, but at an increasing cost, requiring elaborate strategies of "exception barring" and "monster barring." As a mathematics student, I found the story dramatic, even gripping, and when I was done, I had a new view of mathematics, an understanding I wish I'd had before taking all those courses--although I might not have been ready for it then.

Mathematical definitions are, by their nature, as precise as they need to be. As the theory is expanded, the definitions become more specific. And I think the general point of the interplay between proofs and refutations is very relevant to modern mathematics. As an applied

[1] For the *American Mathematical Monthly.* We thank Darren Glass for soliciting this article.

statistician, I'm more of a user than a developer of mathematics--in my career, I've proven only two results that I'd dignify with the title "theorem," and one of them turned out to be false. But the same alternation of proof and counterexample arises for me when developing statistical methods and applying them to live problems.