Hard and Soft Science: Physics vs. Psychology

While doing research for my next book on the differences among so-called hard science, soft science, and pseudoscience, I came across a little-known study published in 1987 by Larry Hedges, then at the Department of Education at the University of Chicago. Hedges was empirically addressing the sort of question that is usually left to a philosopher of science: if it is true—as it is often claimed—that physics (the queen of the hard sciences) "performs" much better than psychology (arguably the Cinderella of the soft sciences), one ought to be able to show, data in hand, that results from physics experiments are "better" than results from psychological experiments. But better in what sense?

Hedges thought that the difference between the two sciences should be evident in the "cumulativeness" of their results: physics should be making progress more steadily and at a faster pace than psychology. This is an important criterion because lack of progress, i.e., lack of cumulative results over time, is one of the distinctive features of pseudoscience. For instance, the idea of intelligent design in biology has made no progress since its last serious articulation by William Paley in 1802. Compare that to the stunning advances in the field of evolutionary biology since Darwin's publication of *On the Origin of Species* in 1859, and one has a good picture of the difference between science and pseudoscience.

As Hedges immediately recognized, however, cumulativeness in science can mean two very distinct, if related, things: on the one hand, one could examine theoretical cumulativeness, i.e., the advancement of a scientific field in terms of how well its theories account for how the world is. For instance, astronomy advanced by a giant leap when it abandoned the Ptolemaic, earth-centered view of the solar system in favor of the Copernican, sun-centered system. Then it made smaller but significant advances by realizing that the planets move in elliptical, not circular orbits; by discovering that the sun is only one star among billions in the Milky Way; and finally by placing our galaxy itself as only one of billions existing in the universe. Judging theoretical cumulativeness, however, is not simple, as it involves a degree of subjectivity, and—more crucially—it requires a long historical perspective. Psychology is a relatively novel science, and it would therefore be rather unfair to compare its theoretical advances with those of modern physics or astronomy, which have

It turns out that the replicability of research findings in psychology (and therefore, presumably, the resulting empirical cumulativeness of that discipline) is no worse (or better) than the replicability of findings in particle physics.

Massimo Pigliucci is professor of evolutionary biology and philosophy at Stony Brook University in New York, a fellow of the American Association for the Advancement of Science, and author of Denying Evolution: Creationism, Scientism and the Nature of Science. *His essays can be found at rationallyspeaking.org.*

had a much longer history.

What then? Hedges opted for a more tractable measure of progress, focusing on empirical cumulativeness. The idea is that if physics, psychology, or any other science is really successful at describing the world as it is, then at a minimum its empirical results (from observation or experiment) ought to be consistent from one publication to another. To put it simply, if Earth is really round with a diameter of about 12,700 kilometers, the different methods to estimate its shape and size ought to yield approximately the same result. If it turns out that some measurement gives us 3,000 kilometers while others go up to as much as 100,000 kilometers, then there is something seriously wrong with the way we do the measurements. Again, the comparison with pseudoscience is obvious: some creationists, for instance, believe that Earth is about 6,000 years old, while others accept the geological figure of four billion years, give or take. This is a staggering discrepancy of five orders of magnitude, which betrays the fact that creationists really have no idea how old Earth is or how to measure it; in turn, this is yet another indication that creationism is no science.

So Hedges went about sifting the literature in particle physics (la crème de la crème in physics research) as well as in a variety of psychological fields, including studies of sex differences, of students' ratings of teaching, of the effect of racial desegregation programs, and others meant to provide a range from quasi-hard psychology (sex differences) to as-soft-as-it-gets research (desegregation studies). He used standard statistical tools to tabulate and compare the results of a variety of studies published over a period of years in several specialized journals.

The results were rather stunning. It turns out that the replicability of research findings in psychology (and therefore, presumably, the resulting empirical cumulativeness of that discipline) is no worse (or better) than the replicability of findings in particle physics. As Hedges put it: "What is surprising is that the research results in the physical sciences are not markedly more consistent than those in the social sciences. The notion that experiments in physics produce strikingly consistent . . . results is simply not supported by the data."

It also turns out that some of the results in physics are much less reliable than one would think. For instance, Hedges compared the data obtained during two series of experiments aimed at estimating the mass of two fundamental particles, the electron and the proton. These are two of the best-known and best-studied particles, so one would expect a high degree of congruency among the outcomes of different experiments. Alas, this was not the case: experiments performed over a period of years (from the early 1960s through the mid-1970s) clearly show that the various estimates were not consistent with each other and their confidence intervals often did not overlap, meaning that the results were significantly different from each other statistically.

Data like these, of course, should not be interpreted as indicating that physicists have no idea what the mass of the electron or the proton is. For one thing, we now have many more experiments, and their results are much more consistent. Moreover, it's not like the estimates reported by Hedges show the kind of huge variation that would make anyone seriously question fundamental aspects of nuclear physics. But the point remains that even the queen of science sometimes gets things wrong over a period of many years, and the quintessential example of soft science, psychology, actually displays a remarkable and surprising degree of consistency in its results. That's something you can quote