# Theory Change and Degrees of Success

## Ludwig Fahrbach[†‡]

Scientific realism is the position that success of a scientific theory licenses an inference to its approximate truth. The argument from pessimistic metainduction maintains that this inference is undermined due to the existence of theories from the history of science that were successful but false. I aim to counter pessimistic metainduction and defend scientific realism. To do this, I adopt a notion of success that admits of degrees and show that our current best theories enjoy far higher degrees of success than any of the successful but refuted theories of the past.

**1. Introduction.** As I understand it here, scientific realism is characterized by adherence to the following inductive principle, which I call the success-to-truth principle: if a theory is successful, then it is approximately true. Examples of successful theories are the atomic theory of matter, the theory of evolution, or the germ theory of disease. As a convention, I generally omit the term "approximately" in describing the consequent of the success-to-truth principle and simply state that successful theories are true. I will also use the term "theory" in a rather generous sense so that it also denotes laws of nature, theoretical statements, sets of theoretical statements, and even classification systems such as the periodic table of elements since I (along with other scientific realists) wish to endorse the truth of these sorts of statements as well.

Realists typically support the success-to-truth principle (or similar principles) with the no-miracles argument (NMA): "Given that a theory enjoys empirical success, wouldn't it be a miracle if it nevertheless were false?" Antirealists of most stripes reject the success-to-truth principle (and similar principles) and whatever arguments realists offer in support of it, for

example, the NMA. Thus, realists and antirealists are engaged in an ongoing dispute over the success-to-truth principle and related principles and the possibility of their justification, in which no side can convince the other side and which goes on without any signs of a resolution.[1]

But now, the antirealist thinks she can offer an *independent* argument—independent of her rejection of the NMA and similar arguments—which undermines the success-to-truth principle. This argument is the pessimistic metainduction (PMI). The PMI moves from the premise that the history of science is full of theories that were once successful and accepted by scientists as true but later refuted and abandoned. Let's assume for the time being that this premise is correct. Then these successful but false theories constitute *counterinstances* to the inference from success to truth. In other words, the success-to-truth principle has had a really bad track record, which counts strongly against its being valid.[2]

The premise of the PMI about the widespread occurrence of successful but false theories in the history of science requires evidence. Thus, Laudan (1981) famously presents the following list of examples of such theories:

- the crystalline spheres of ancient and medieval astronomy
- the humoral theory of medicine
- the effluvial theory of static electricity
- 'catastrophist' geology (including Noah's deluge)
- the phlogiston theory of chemistry
- the caloric theory of heat
- the vibratory theory of heat
- the vital force theories of physiology
- electromagnetic ether
- optical ether
- the theory of circular inertia
- theories of spontaneous generation

The antirealist then argues that, even if judged from the perspective of the realist, that is, starting from the confirmational views of the realist (the success-to-truth principle and the NMA) and disregarding the confirmational views of the antirealist, the success-to-truth principle has to be given up. From this perspective two arguments concerning that principle have to be considered, the NMA and the PMI. The NMA supports it; the PMI undermines it. The two arguments have to be balanced against

1. For recent book-length treatments of the scientific realism debate, see Leplin (1997), Kukla (1998), Psillos (1999), Stanford (2006), Chakravartty (2007), Turner (2007), and Ladyman and Ross (2008).

2. Compare, among others, Psillos (1999, chap. 5), Kitcher (2001), Lewis (2001), Saatsi (2004), Devitt (2005), and Schurz (2009).

each other. The antirealist maintains that the result of the balancing is that the PMI is much stronger than the NMA. Whereas the NMA seems to be a priori and ultimately based on intuitions, the premise of the PMI is based on empirical evidence from the history of science and provides many concrete counterexamples against the inference from success to truth. What better case against an inference than counterexamples can you provide? Hence, the PMI trumps the NMA. The antirealist concludes that even if one starts from endorsing the realist's confirmational views, one has to change one's view about the success-to-truth principle and admit that it is undermined by the past of science.

My goal in this article is to outline a novel defense of scientific realism against the attack of the PMI. To do so I will use a graded notion of success of scientific theories. I will compare our current best theories and the refuted theories of the past with respect to degrees of success. The result of the comparison is the main thesis of this article: our current best theories enjoy far higher degrees of success than any of the refuted theories of the past, which enjoyed only fairly low degrees of success.

My goal is modest in that I will not engage in the realism debate as it was briefly presented above but will argue entirely from the perspective of the realist, in the same way as the antirealist just did, namely, relying exclusively on the realists' confirmational views (the success-to-truth principle and the NMA) and disregarding the confirmational views of antirealists. Hence, my goal is solely defensive. I want to show that by suitably modifying the realist position, it can be shown not to be in conflict with the history of science; in particular, by suitably modifying the success-to-truth principle it can be saved from the counterexamples from the history of science.

**2. Degrees of Success.** The notion of success as it appears in the literature is fairly vague. I will make it more precise in two ways. First, I will connect it with standard ideas of how theories are tested by observation. Second, I will assume that it admits of degrees. Consider the following standard account of theory testing. In order to test a theory, scientists derive predictions from the theory and make observations. A particular test of a theory consists in comparing some prediction of the theory with some observation. If prediction and observation agree, the theory enjoys some measure of success. If prediction and observation do not agree, the theory suffers from an anomaly. As long as the anomaly is not significant or the anomalies do not accumulate, they do not refute the theory. If the anomaly is significant or the anomalies do accumulate, the theory counts as refuted. Of course, this account of theory testing and empirical success is rather minimal in several respects and could be made more precise in many ways, but it is all we will need here.

I will use the notion of a test in a rather broad sense here to cover all

cases in which scientists are aware of the theory coming into contact with experience somehow so that it is possible for the theory to fail. I also use the term "prediction of a theory" in a rather broad sense to denote any observable consequence of the theory scientists are aware of. It will later be important that different tests can differ with respect to their quality, for example, with respect to the precision of the data and the precision of the predictions involved. It will also be important that the overall degree of success of a theory at a given time is in part determined by the total number and diversity of all the tests that the theory has passed until that time.

I will now aim to compare the degrees of success of the refuted theories with those of our current best theories. What follows is a list meant to be representative of our current best theories (remember that the realist endorses the *approximate* truth of those theories):

- the periodic table of elements
- the theory of evolution[3]
- "Stars are like our sun."
- the conservation of mass energy
- the germ theory of disease
- the kinetic gas theory
- "All organisms on Earth consist of cells."
- $E = mc^2$
- And so on

In order to compare the degrees of success of current best theories with the successful but refuted theories of the past, I want to employ five indicators of success. I call them "indicators of success" because, as I argue in the last section, they are positively correlated with the degrees of success of theories. The first indicator is the amount of scientific work done by scientists until some time. The second, third, and fourth indicators are the amount, diversity, and precision of scientific data and observations gathered by scientists until some time. The fifth indicator is the amount of computing power available to scientists at some time. I will examine the growth of these indicators over the history of science. From their growth I will infer (in the last section) the main thesis of this article according to which our current best theories enjoy far higher degrees of success than the refuted theories of the past.

**3. Scientific Work.** The first indicator of success is the amount of scientific work done by scientists in some period of time, where "scientific work"

3. The theory of evolution may be taken to consist of two claims: that all organisms on earth are related by common ancestors and that natural selection is an important force for change.
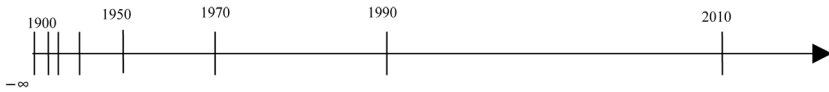
Figure 1. Timeline weighted in such a way that the length of any interval is proportional to the amount of scientific work done in that interval.

means such things as making observations, performing experiments, testing theories, and so on. It can plausibly be measured with the help of two quantities: the number of journal articles published in the respective period of time and the number of scientists working in the respective period of time. Both quantities show roughly the same kind of growth over the last 300 years: the number of journal articles published by scientists every year has doubled every 15–20 years over the last 300 years, and roughly the same holds, as far as we know, for the number of scientists. This implies, for example, that three-quarters of all scientific work ever done was done in the last 30–40 years, while one-quarter was done in all the time before (see fig. 1).

A telling example is the case of chemistry and the periodic table of elements. As in the rest of science, manpower and publications in chemistry have risen exponentially. Schummer (1999) shows that during the past 200 years the growth of scientific work in chemistry has meant that the number of newly discovered or produced chemical substances has risen exponentially with a fairly constant doubling rate of 13 years. The periodic table of elements implies constraints about, for example, the features of every chemical substance and chemical reaction. Hence, every newly discovered or produced chemical substance provides an occasion for a (mostly not very severe) test of the periodic table of elements. That the periodic table of elements has been entirely stable for many decades shows that the tests have always been passed. Mostly they have only provided a weak increase in degrees of success. But because the number of such tests has been huge, the overall increase in degrees of success of the periodic table of elements has been huge.

**4. Amount and Diversity of Data.** The second indicator of success is the amount of data gathered by scientists until some time. Here, we observe that for many kinds of data the amount has grown at a very high rate. First, in some disciplines such as paleontology or chemistry, it is often still the scientists themselves who gather or produce the data, for example, searching for fossils or synthesizing chemical substances. For such disciplines, it is plausible that the amount of data gathered or produced by scientists has grown very roughly proportionally to the number of sci-
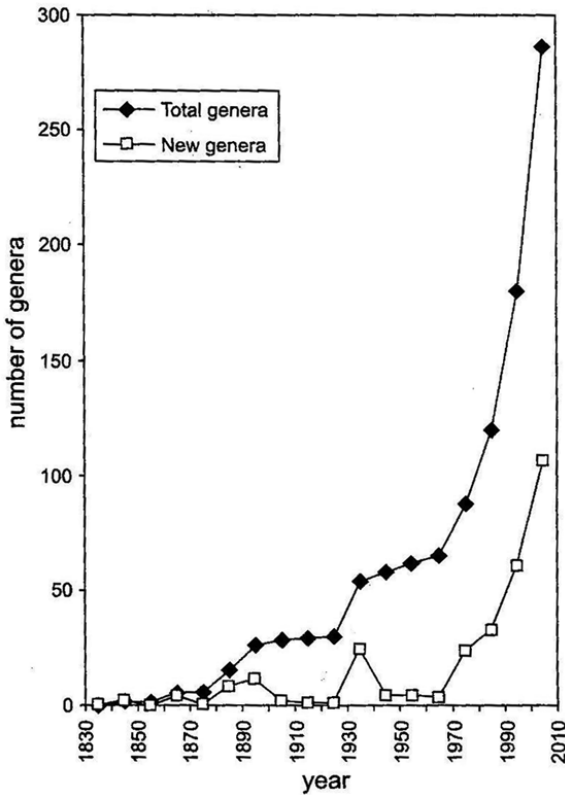
Figure 2. Growth of the fossil record of mammals from the age of dinosaurs (245–66 million years ago). From Kielan-Jaworowska, Cifelli, and Luo (2004), S.7.

entists in the field. Today there are at least 30 times as many scientists as in 1900 and at least 1,000 times as many scientists as in 1800. Therefore, in such disciplines the amount of data has often risen in a similar fashion. For example, figure 2 depicts the increasing rate at which archaeologists have uncovered new fossils of Triassic era mammals. Such increases are entirely typical for the growth of the fossil record.

Second and more important, for many kinds of data the amount of their growth has been far greater than the growth of scientific manpower due to better instruments and computer technology. In many disciplines, data are nowadays gathered automatically (cf. Humphreys 2004, 6–8). For example, during the last 6 years the Sloan Digital Sky Survey has "measured [the] precise brightnesses and positions for hundreds of millions of galaxies, stars and quasars . . . [mapping] in detail one-quarter of the

entire sky" (http://www.sdss.org/; see also Kennicutt 2007). By comparison, the most ambitious such project at the beginning of the twentieth century, a survey of the sky conducted at Harvard and completed in 1908, measured and cataloged the brightnesses and positions of 45,000 stars.

Another example is provided by the sequencing of DNA. Here, over the last 20 years the overall number of decoded DNA sequences has grown with a fairly stable doubling rate of 18 months (National Center for Biotechnology Information, GenBank Statistics, http://www.ncbi.nlm.nih .gov/Genbank/genbankstats.html). This means a growth by a factor of 100 every 10 years from 1984 (around 4,000 sequences) to 2004 (around 40 million sequences). Again, this is not possible without automation. In these and many other fields, the automatic gathering of data has lead to truly gigantic amounts of data.

Note that the data sets presented above are not such that all the pieces of data are of the same narrow kind. Instead they exhibit a high diversity. For example, it is not the case that paleontologists examine the same features of the same fossils again and again or that all fossils they find are of the same species and the same age. Instead, paleontologists naturally look for fossils from different locations and strata, and what they find are mostly different species. Thus, in figure 2, the $y$-axis depicts the number of genera of mammalian fossils.[4] Similarly, it is not the case that chemists examine the same features of the same substance again and again; instead they create, as we saw, new substances incessantly. The millions of chemical substances that have been synthesized so far clearly represent an extremely high variety of evidence.

**5. Precision of Data and Computing Power.** The fourth indicator of success is the precision of data. Data become more precise when scientists improve already existing kinds of instruments and measurement techniques or develop new kinds of instruments and measurement techniques. This happens all the time, of course, and has lead to constant improvement in the precision of data over the last few centuries and especially the last few decades. Often, the improvements were by great leaps. Examples abound. Let me just mention two especially interesting ones.

The first example concerns the measurement of distances between places on the surface of the earth for the purpose of determining details about the movement of tectonic plates. The kinds of measuring techniques available before the 1980s required years to produce meaningful data. In the

---

4. The fossil record as evidence for the theory of evolution has only a limited diversity. However, fossils are just one kind of evidence for the theory of evolution among very many very different kinds of evidence from very different sources such as biogeography, genetics, embryology, molecular biology, etc.

1980s this changed dramatically through the advent of global positioning systems (GPS). The precision increased a thousandfold. In consequence, determining the movements of tectonic plates became rather easy and very reliable. In addition to the substantial increase in the precision of the data, the data also exhibit a large variety, originating from thousands of GPS stations from many different places on the earth. The second example is the increase in precision of time measurement. Since the 1950s, the precision of atomic clocks has increased by at least one digit per decade (Sullivan 2001, 6). Today the best clocks, so-called optical clocks, maintain an accuracy $10^{-12}$ seconds per day. Needless to say, precise measurements of time are vital in very many different scientific fields (e.g., for GPS) and have led to a high increase in success for many of our current best theories over the last few decades.

The fifth indicator of success is computing power. Until 50 years ago computations were done by humans; therefore, overall human computing power rose at least at the rate of the number of scientists and actually considerably more due to the introduction of instruments like logarithmic tables and slide rules. Furthermore, in the last 50 years computing power of digital computers doubled roughly every 2 years. Such a growth is, of course, much greater than the growth of the number of scientists and journal articles. The increase in computing power is then connected with increases in success of theories in straightforward ways. For example, with more computing power, scientists can solve more equations and can solve them quicker. This results in a higher number of predictions from theories, which can thereby exhibit more diversity. It is then highly plausible that the growth of computing power has contributed strongly to the increase in degrees of success of our current best theories. Thus, Humphreys remarks that "much of the success of the modern physical sciences is due to calculation" (2004, 55, esp. chap. 3).

**6. Saving Realism.** We have seen that several indicators of success of scientific theories have enjoyed an enormous increase over the last few decades. Previously, almost all of them were quite low; today, they are all very high. From this we can infer the degrees of success of both our current best theories and the refuted theories of the past, thereby arriving at the main thesis of the article: our current best theories enjoy far higher degrees of success than any of the successful but refuted theories of the past, which enjoyed only quite modest degrees of success.

The argument proceeds along these lines. On the one hand, our current best theories were around in the recent past (the last 50–80 years, say) and have therefore profited from the enormous increase in the indicators of success in that time: the amount, diversity, and precision of data have increased enormously, and the same holds for the computing power of

scientists. Therefore, our current best theories have been subject to an enormous number of tests (even if many or most of them have been only of a weaker kind). Hence, they have received a big boost in their degrees of success in the recent past, an increase that is far greater than any increase in success of any theories of earlier times. This inference proceeds on a general level, but we could also see directly for a number of specific theories, how their degrees of success profited from the increase of the indicators. On the other hand, all the theories on Laudan's list are rather old, namely, more than 100 years old. The same holds for practically all examples of theory changes offered in the philosophical literature.[5] At those times practically all indicators were quite low: the data sets were comparatively small, showed comparatively little diversity and precision, and given their computing power scientists could only produce comparatively few and imprecise predictions from the theories. Therefore, the theories of those times were subject to comparatively few tests that had comparatively modest severity. It follows that the degrees of success of theories of those times, in particular the degrees of success of the refuted theories, were quite modest. Putting both arguments together, we arrive at the main thesis of the article.

Of course, the considerations of this article are no more than an outline of an argument that needs to be refined in several respects. Also, a number of objections have to be dealt with. (Did some theories not enjoy very high degrees of success already early on in the history of science? Should the occurrence of refuted theories not be extrapolated from lower degrees of success to higher ones, thereby undermining the success-to-truth principle? What about refuted theories from the recent past? And so on.) I think all objections can be met, but I cannot discuss them here. I will simply conclude by noting how the central argument of the article, if successful, saves realism.

Realism is saved by modifying the success-to-truth principle in the following way: if a theory enjoys *very high* degrees of success, it is approximately true. Because the modified success-to-truth principle is not threatened by counterexamples, the realist can modify his position so that it consists in the endorsement of the modified success-to-truth principle. This modified form of realism is saved from the PMI. It is a version of realism that is compatible with the history of science. The realist can then support the modified success-to-truth principle with the NMA. Finally, he can apply the modified success-to-truth principle to our current best theories and infer that they are approximately true.

---

5. Some important theory changes occurred in fundamental physics in the twentieth century, but I think that theories from fundamental physics represent a special case with respect to the success-to-truth inference and should be examined separately.

REFERENCES

Chakravartty, Anjan. 2007. *A Metaphysics for Scientific Realism: Knowing the Unobservable*. Cambridge: Cambridge University Press.

Devitt, Michael. 2005. "Scientific Realism." In *The Oxford Handbook of Contemporary Analytic Philosophy*, ed. Frank Jackson and Michael Smith, 767–91. Oxford: Oxford University Press.

Humphreys, Paul. 2004. *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. Oxford: Oxford University Press.

Kennicutt, Robert C., Jr. 2007. "Sloan at Five." *Nature* 450 (November 22): 488–89.

Kielan-Jaworowska, Zofia, Richard Cifelli, and Zhe-Xi Luo. 2004. *Mammals from the Age of Dinosaurs: Origins, Evolution, and Structure*. New York: Columbia University Press.

Kitcher, Philip. 2001. "Real Realism: The Galilean Strategy." *Philosophical Review* 110 (2): 151–97.

Kukla, André. 1998. *Studies in Scientific Realism*. Oxford: Oxford University Press.

Ladyman, James, and Don Ross. 2008. *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Laudan, Larry. 1981. "A Refutation of Convergent Realism." *Philosophy of Science* 48 (March): 19–49.

Leplin, Jarrett. 1997. *A Novel Defence of Scientific Realism*. Oxford: Oxford University Press.

Lewis, Peter. 2001. "Why the Pessimistic Induction Is a Fallacy." *Synthese* 129:371–80.

Psillos, Stathis. 1999. *Scientific Realism: How Science Tracks Truth*. New York: Routledge.

Saatsi, Juha. 2004. "On the Pessimistic Induction and Two Fallacies." *Philosophy of Science* 72 (Proceedings): 1088–98.

Schummer, Jochen. 1999. "Coping with the Growth of Chemical Knowledge." *Educación Química* 10 (2): 92–101.

Schurz, Gerhard. 2009. "When Empirical Success Implies Theoretical Reference: A Structural Correspondence Theorem." *British Journal for the Philosophy of Science* 60 (1): 101–33.

Stanford, P. Kyle. 2006. *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford: Oxford University Press.

Sullivan, D. B. 2001. "Time and Frequency Measurement at NIST: The First 100 Years." In *Proceedings of the 2001 IEEE International Frequency Control Symposium and PDA Exhibition*, 4–7. Piscataway, NJ: IEEE.

Turner, Derek. 2007. *Making Prehistory: Historical Science and the Scientific Realism Debate*. Cambridge: Cambridge University Press.