

Pacific Standard

ENVIRONMENT

RESEARCHERS TURN TO BIG DATA TO JUSTIFY BASIC SCIENCE

A new method of citation network analysis could be used to make the case for the importance of basic research.

MICHAEL WHITE · NOV 9, 2015

For more than half a century, most basic scientific research conducted in the United States has been paid for by the federal government. This major investment in basic science is based on a key premise: New technologies that solve practical problems and help us lead healthier, more productive lives emerge from continued discoveries in fundamental science. Without advances in basic physics, chemistry, molecular biology, and computer science, we wouldn't have smartphones, hybrid cars, Lipitor, or search engines. "Basic research is the pacemaker of technological progress," wrote Vannevar Bush, director of the government's research efforts during World War II and the architect of our modern approach to government-funded research. This idea is widely accepted by scientists. But is it true? How important is basic research, really?

Scientists tend to defend basic research in a rather unscientific way—by telling anecdotes. Many of these anecdotes are persuasive, and it's clearly true that some discoveries in esoteric fields of basic research have paved the way for the technologies that we now depend on every day. Without Einstein's abstract theories of relativity, there would be no GPS for our smartphones to tap into. If scientists hadn't studied the genetics of bacteria in the 1940s, we wouldn't know what a gene is—and modern medicine as we know it wouldn't exist. Rather than offering a systematic evaluation of the contribution of basic science, basic researchers rely heavily on handpicked examples of success to justify the government funding on which their work depends.

THE GOVERNMENT'S INVESTMENT IN BIOMEDICAL RESEARCH HAS DECLINED BY 22 PERCENT SINCE 2003.

For example, in 2013 a major scientific society sponsored a competition called "Stand Up for Science," in which contestants had to make a short video pitch to the American taxpayer "aimed at increasing awareness of federal funding support for biological and biomedical science." The winning entry, by a group of graduate students at the University of California–San Francisco, was yet another anecdote. In the video, the students asked, "The year is 1960: You have \$10 to spend on research. Would you spend it on developing an affordable treatment for diabetes, or conducting basic research on how bacteria protect themselves?" With hindsight, the answer is clear: Basic research on how bacteria protect themselves rapidly led to one of the most revolutionary advances in the treatment of diabetes ever—genetically engineered bacteria that produce human insulin, a drug used by millions of diabetics worldwide.

It's easy to take note of a handful of spectacular successes, and anecdotes are much more likely to excite our imaginations than a more rigorous set of statistical metrics. But do we really know how well our overall investment in basic research pays off? This is an incredibly difficult question to answer because the path from basic research to new technologies is typically indirect, and usually takes decades to traverse. Einstein's theory of relativity was essential for the development of GPS, but many other scientific and technological advances also had to be in place before it could become feasible. Because the development of a single technology often occurs at the confluence of many different lines of research, it's hard to predict how basic research will pay off in the future.

Even assessing how research has paid off in the past is difficult. In a 2014 report on the impact of basic research, a committee of the National Academy of Sciences noted that, "It is virtually impossible to extrapolate the impact of a single research program forward through multiple levels of development and commercialization because of the resulting technology's combination with other technologies to make an eventual impact on economic growth or some societal goal."

So does that mean that we can't rigorously and quantitatively justify our government's long-term investment in basic research? Not necessarily. In September, a team of scientists at UCSF published a new method for systematically assessing how basic research contributes to concrete medical advances that have a tangible impact on people's lives. The researchers turned to a set of now-common computational techniques, data mining and network analysis, to trace the non-linear path from research to new clinical treatments. It works like this: Beginning with the clinical trial documents for a drug recently approved by the Food and Drug Administration, the method uses online databases to trace a network of citations. By mapping successive rounds of citations—the research papers cited by the research papers that are cited, in turn, by the clinical trial documents—one can build a "cure discovery citation network" of research papers, scientists, and institutions "that were most important in establishing the base of knowledge that enabled the successful drug development program."

The scientists applied their method to a major, newly approved cystic fibrosis drug, Ivacaftor. They found that the research network behind Ivacaftor was enormous: Almost 3,000 scientists—whose research spanned 59 years—laid the groundwork for this drug. But this might be an overestimate; many of these scientists might simply have been cited once in the network, and thus may have made only a marginal contribution to the development of Ivacaftor. To look at the *major* contributions, the researchers narrowed in on scientists and research papers that were most highly connected in the network, the "elite performers" who made a sustained contribution to the science behind the new drug. That network was much smaller than the original one, but still surprisingly large: 33 scientists, who published 355 relevant research papers over 47 years. No wonder the impact of basic research on new technologies and treatments is so hard to grasp: The path from research to drug spans decades and involves dozens of researchers and hundreds of studies.

Citation networks aren't a perfect way to assess the value of basic research—for example, they don't address the question of how much basic research *doesn't* lead to practical payoffs. But this kind of network analysis does show, rigorously and dramatically, what scientists' favorite anecdotes only hint at: Technological and medical progress is built on a foundation of science that is both broad and deep. As the authors of the study argue, their results demonstrate "that future cures will depend on broadly based public support of life sciences." Unfortunately, that support has been steadily eroding for more than a decade, as the government's investment in biomedical research has declined by 22 percent since 2003. If one of the major goals of our society is to improve quality of life through progress in technology and medicine, then, as the study of cure networks confirms, the failure to invest in a foundation of basic research will ultimately be self-defeating.

Inside the Lab explores the promise and hype of genetics research and advancements in medicine.

RELATED



EDUCATION
**CAMPUS RESEARCH
BACK TO BASICS**

ENVIRONMENT
**HOW BIG DATA IS
CHANGING SCIENCE
(AND SOCIETY)**



ENVIRONMENT
**WHEN BIG DATA
BECOMES BAD
DATA**



ENVIRONMENT
**DO WE NEED
FORMAL QUALITY
STANDARDS FOR
SCIENCE?**



ENVIRONMENT
**WHY CURIOSITY
SHOULD DRIVE
SCIENTIFIC AG**