What’s So Special About Science (And How Much Should We Spend on It?)

William H. Press

Scientific research probes the deepest mysteries of the universe and of living things, and it creates applications and technologies that benefit humanity and create wealth. This “Beauty and Benefits of Science” is the theme of this 2013 AAAS Annual Meeting.

The subject of my address is a different kind of mystery, although it is also related to this theme. It is the mystery of why society is willing to support an endeavor as abstract and altruistic as basic scientific research and an enterprise as large and practical as the research and development (R&D) enterprise as a whole. Put differently, it is the mystery that a unified scientific enterprise can be simultaneously the seed corn for economic advance and the confectionary corn syrup of pure, curiosity-driven scientific discovery.

The view that science can be supported as a contribution to the intellectual richness of the world has a distinguished list of adherents. In 1969, Robert Wilson explained what Fermilab would do for the country by saying, “It has nothing to do directly with defending our country except to make it worth defending” (1). And, almost two centuries earlier, in his first annual address to Congress, George Washington wrote, “[t]here is nothing which can better deserve your patronage, than the promotion of Science and Literature. Knowledge is in every country the surest basis of publick happiness” (2).

Indeed, U.S. taxpayers are, to some extent, willing to pay for activities that enrich American social and cultural capital without having a direct economic benefit. Congress, up to now, has appropriated about $150 million a year for the National Endowment for the Humanities (NEH) (3). However, by contrast, Congress appropriates about $40 billion a year for basic research (4). If you plot a bar graph with these three numbers, you can barely see that the NEA and NEH numbers are not zero.

It is evident that society is willing to pay much more for curiosity-driven research in science than for the analogous thought- and beauty-driven practice of the arts and humanities. It is easy to guess the reason: the link, sometimes subtle but repeatedly established over time, between investment in basic research and macroeconomic growth. Discovery leads to technology and invention, which lead to new products, jobs, and industries.

Such is the case that we scientists need to reinforce in the austere times that we face. However, mere repetition is not an effective strategy. In today’s lean times, we need to articulate our case more powerfully and in a more sophisticated way than in more prosperous times. A skeptical and stressed Congress is entitled to wonder whether scientists are the geese that lay golden eggs or just another group of pigs at the trough.

More Than a Century of Exponential Growth

Figure 1 shows the growth in U.S. gross domestic product (GDP) per capita over the past 130 years. If we ignore a few bumps with time scales of a decade or so, the curve is surprisingly well fit by a pure exponential. Note that the curve is not plotting GDP, which would grow with population and the overall size of the U.S. economy, but GDP per capita, which reflects something like the average income of each individual. Ameri-
can individual income has grown exponentially. (That is on average. The distribution of that income within the population is another matter entirely.)

In the wake of the financial crisis, economist Robert Gordon advanced the theory that 2008 was the beginning of a new normal where exponential growth could no longer be assumed (5). However, one sees plenty of other temporary blips in the historical record, the Great Depression of the 1930s being most visible. Any of these fluctuations might have been mistakenly overinterpreted. Indeed, from a low in 2009, GDP per capita has increased in each of the past 3 years (6).

Turn now to Fig. 2, from the Battelle Memorial Institute and R&D Magazine (7). The horizontal axis shows the fraction of their GDP that countries of the world spend on research and development. The vertical axis shows what fraction of their populations are scientists or engineers. Unsurprisingly, these are highly correlated. The area of each country’s circle is its total R&D spending. The United States is the largest circle, just short of 3% of GDP, and at about a half percent scientists and engineers in the population. Notice that the United States is not the most intensive investor, neither as a percentage of GDP nor as a fraction of its population.

It is striking that the powerhouse industrial nations of the past 50 years (e.g., United States, Germany, Japan, and South Korea), lead the pack at the upper right, together with the Scandinavian countries and (recently) Singapore. These countries might thus be identified as the technology leaders. The United States ranks third or fourth among large economies in what we spend per capita on R&D and ranks eighth among all countries (8).

A second cluster of developed nations spends about a third less and has correspondingly fewer scientists and engineers: United Kingdom, France, Canada, Australia, and Russia, among others. These industrialized countries might thus be identified as the technology followers. A third cluster at the lower left comprises the developing countries of the world, China being by far the largest.
It would be overly simplistic to ask, “Which of these three clubs do you want to be in?” and then equate spending on research with the sustained economic growth that gives the powerhouse industrial nations their high standard of living. Correlation, we know, is not the same as causality. Are we richer because we spend more on R&D? Or do we spend more on R&D because we are richer, and can afford to do so? Or is there some other factor, causative of both?

The Solow Residual Is Technology

Luckily, we know the answer, thanks to intensive study by economists, starting with Robert Solow (who won a Nobel Prize for this work) (9), as well as luminaries such as Ken Arrow, Zvi Griliches, and many others. Starting after World War II and continuing in the 1970s and 1980s, economists have largely settled the question of where the exponential growth of the U.S. economy has come from and what the factors that produced it were.

Although I am not an economist, I want to give you a simple version of the story, because it is one that I think scientists need to understand as part of making the case for what we do. Exponential growth comes from positive feedback, where the production of something enables you to produce even more. In other words, during the exponential growth that we have experienced, something produced must itself have been a factor of production. Economists have been studying factors of production since the time of Adam Smith and David Ricardo more than 200 years ago. The classic factors of production are land, labor, and capital. Leaping over two centuries of economic history, we might modernize these by adding various forms of human capital, such as investment in education. We might also add intellectual property as capital, which today is no less important than factories and machinery. Interestingly, this change was formally made by the U.S. Bureau of Economic Analysis only this year, resulting in a $560 billion “increase” in the U.S. GDP (10).

A relatively fixed factor of production such as land (today we would include all natural resources) cannot produce positive feedback for exponential growth. Land does not multiply to produce more land. Other natural resources are at best augmented by one-time discoveries; at worst, they are depleted. Labor can multiply to produce more labor, in the sense of a GDP that increases in proportion to population growth. But growth of labor cannot by itself produce more income per person. Therefore, it is uniquely what economists call capital (albeit in its broadest sense, including human, intellectual, and environmental capital) that can fuel exponential growth of the economy.

Solow and subsequent researchers studied this and attempted a detailed accounting of how much growth was due to each of the known factors of production. Their finding was that at most half of the historical growth could be explained by known factors. The unexplained part, sometimes estimated as large as 85%, was termed the Solow residual. Subsequent work showed that the bulk of the Solow residual could in fact be explained by positing a new factor of production: technological progress. Technology is not exactly a form of capital (although it is related to human and intellectual capital), because the technology state of the art is not exactly owned by anyone. Yet it is capable of generating positive feedback. As a factor of production, technology produces wealth and produces more technological progress, enabling a virtuous cycle of exponential growth (11, 12).

Once this new factor was identified, many studies from the 1970s forward have taken small, studiable slices of the economy and tried to estimate the annual rates of return on investment (ROI) in basic research—the fuel for technological progress. As AAAS members, we may all be gratified that these studies nearly all show high rates of return. Many institutions, including our universities and retirement funds, accept 5% sustained ROI as a decent return. Yet investments in basic research are variously estimated as ultimately returning between 20% and 60% per year (13).

The Appropriability Conundrum

If that were the whole story, my advice today would be simple enough: Withdraw all your retirement funds and invest them in your laboratory! But, of course, it’s not the whole story. What makes this advice bad (apart from its being a risky undiversified investment) is another important concept from neoclassical economics, appropriability. Appropriability means the following: How well do the rewards flow back to the investor who actually takes the risk and puts up the money?

The returns from basic research are large, but they are not very appropriate. The nature of basic research is that its results flow to the rest of the world. Although basic research can be turned into applied research—into patents, products, and eventually economic growth—this may not necessarily occur in the laboratory where the work is originally done or even in the same country. So the appropriability of basic research is low. The investor generally does not get enough of the reward.

Basic research leading to scientific discovery is thus a public good. It will benefit all. But,
ASSOCIATION AFFAIRS

because the private incentive to pay for basic research is therefore attenuated, the private sector as a matter of economic self-interest is likely to underinvest in it. Therefore, if the full social potential of this public good is to be realized, the investment must come from government. Indeed, government in the United States does support the lion's share of the nation's basic research endeavors.

Figure 3, based on data from the National Science Foundation (NSF) (14), shows total U.S. support of R&D over the past 50 years as a percentage of GDP and how it is divided between federal and nonfederal (dominantly industry) funding. Looking now at all R&D (not just basic research), one sees that, post-Sputnik, total R&D per GDP has fluctuated between about 2.2 and 2.8%. Without minimizing the importance of the up-and-down fluctuations (these are, respectively, beneficial or destructive to the research enterprise), U.S. investment in R&D has largely kept pace with the size of the U.S. economy. Indeed, following Solow's theory, it is likely responsible for up to 85% of U.S. economic growth.

What has changed dramatically over 50 years is the fraction of total R&D support from the federal government. In the years immediately after Sputnik, about two-thirds of total R&D was federally funded, whereas one-third was funded by the private sector. Today the proportions are almost exactly reversed: About two-thirds of U.S. R&D is paid for by industry.

This change drives a shift toward appropriable R&D, that is, more “D” and less “R,” because that is the kind of investment that more likely yields products and services that can get to the market quickly, thus yielding returns for the investors who invest in the companies that fund the work. The federally funded fraction, which has been decreasing, is the less appropriable part, the part that is a public good—the seed corn.

What has driven this trend? (I am again simplifying a complicated piece of economic history.) In 1945, just after World War II, the United States accounted for 50% of the total world GDP. Therefore, federal investment in R&D was almost guaranteed to be at least 50% appropriable, with its returns accruing to the U.S. economy. In fact, the appropriability was likely even larger, because, until as late as the 1960s, America's potential competitors, Japan and Europe, lacked the full capacity to absorb the results of U.S.-funded basic research. By contrast, information flows today are so rapid that anyone, anywhere, can potentially be the entrepreneur who recognizes the economic potential of scientific discoveries. U.S. scientific results used to be exported at the speed of ocean mail or international student Ph.D. programs; today, they are exported at Internet speeds.

Fast forward to the 1990s. The industrialized world is a much more competitive place. In the United States, the era of regulated monopolies (such as the Bell System, which paid for the hugely productive Bell Laboratories) is over. Large multinational corporations are becoming more dominant. Quarter-by-quarter accountability to shareholders is now an influential pressure point. Increased industry investment in R&D is a rational response, because R&D is (as discussed) a good investment. However, industry investment in nonappropriable R&D, such as basic research, is becoming harder to justify, both because of its longer time horizon and because of the risk that others will appropriate the results.

Indeed, the data from 1990 to 2007 support this interpretation. During this whole period, industry investment in basic research stagnated at about $10 billion (constant 2005 dollars). Thanks to increased federal and other funding during most of this period, the total amount of basic research did increase (even faster than the overall size of the economy). However, only in 2008 to 2010 in the wake of the financial crisis, did industry increase its spending on basic research, and this may be simply an artifact of renewed attention by industry to R&D generally in that period. In fact, industry spending on basic research again decreased in 2011. Figure 4 shows the data for total annual investment in U.S. basic research from 1965 to 2011 and how it has been divided among industry, the federal government, and other funding sources.

Overgrazing the Commons of Basic Research

The stagnation of industry's funding of basic research in the period from 1990 to 2007 presages an even more alarming trend, also shown in Fig. 4: Total federal spending on basic research has stagnated since 2003 at about $37 billion (2005 dollars). This creates an extremely dangerous situation, and for two related reasons.

First, 10 years into this trend, the federal budget is now facing a time of unprecedented austerity. Basic research needs to grow at least as fast as the economy because it is part of the feedback loop for most of the economy's growth. Industry support, even if it proves to be on the rise, supplies only a small fraction of what is needed; therefore, federal funding must make up the bulk of the difference. In today's environment of austerity, the rate of increase in federal support needed to make up for a decade of neglect, and to keep pace with the recovering economy, will be a very tough sell.

Second, the traditional bipartisan support enjoyed by federal spending on basic research has always been conditioned on its being a public good whose benefit accrues to the U.S. taxpayer. Therefore, any decrease in that investment's appropriability—now on the
scale of nations, not individual corporations—
decreases the attractiveness of the investment.
In other words, the danger of the 2010s is that
nations such as the United States will make
the same decision that individual corporations
made in the 1990s and invest relatively less
in basic research. In effect, the federal gov-
ernment already made this decision around
2003, because its support of basic research
has stagnated in constant dollars since then. If
other countries follow suit, this could lead to
what we have never seen before in 150 years
of exponential growth: a zero-sum world of
research. This would be a world in which
every country tries to gain the benefit of what-
ever basic research is funded by others, but
no country is willing to make investments them-
selves. Rather, all scramble to compete over
diminishing returns from past investments.
This scenario is what economists call the trag-
ey of the commons. No one will invest to main-
tain something that ben-
fits all, given the option
of instead being a free
rider. The classic trag-
ey of the commons is a piece of grassland
owned in common. Everyone is entitled to
graze their sheep, and no one pays for the
land’s upkeep. And, after a while, there are a
lot of hungry sheep and no grass.
Basic research is a commons. It can be
overgrazed and underreplenished. There is
no Global Science Foundation that has an
annual appropriation of the world’s taxes to
invest for the ultimate good of everyone in
the world. Yet, not all nations can be free riders,
especially not the United States.
The current situation is dangerous. Short-
term actions in a time of budget crisis and
financial austerity might become the triggers
of long-term underinvestment in the ultimate
fuel of economic growth, basic research in
science. Can this tragedy be avoided? Pos-
sibly. Let me sketch two paths forward
for avoiding such a negative outcome.

**Geographical Anchors and Nonmonetizable Benefits**
The first path forward is to adopt policies
that maintain the appropriability of returns
from U.S. basic research to the U.S. econ-
omy. This requires finesse because policies
that conflict with the fundamental open-
ness of such research are sure to be coun-
terproductive. Yet, even in the Internet age,
there are strategies that can geographically
anchor the transformation of research dis-
scoveries into economic growth. One such
phenomenon is the growth of research hubs
around America’s great research universities.
The older examples are Silicon Valley on the
West Coast and Route 128 around Boston,
but examples now include established and
growing high-tech corridors in or near New York (Silicon Alley), Wash-
ton (Dulles Technology Corridor), Chicago
(Golden Corridor), Portland (Silicon For-
est), Raleigh-Durham (Research Triangle),
and other places.

The point is that research universities are
geographically rooted communities,
and they are not easily
exported. One can
imagine a corporation
moving its R&D overseas,
but it is difficult
to imagine that Stan-
ford or the Massachu-
setts Institute of Tech-
ology will decide to
up and relocate. Prox-
imity and day-to-day
human social interactions breed more effi-
cient technology transfer. Thus, research uni-
versities and the industry that they engen-
der serve to increase the appropriable return
for the nations and regions that host them. A
recent report (15) by the President’s Counci-
il of Advisors on Science and Technology
(PCAST), in which I participated, delves fur-
ther into the new importance of the out-
standing U.S. research universities. That report
emphasizes their new dual mission, to pro-
tect and advance basic research and to make
that basic research more easily translatable
into application and technology in a way that
remains geographically rooted.

There are also other kinds of rooted infra-
structure, such as research parks, research
incubators, national user facilities (such as
synchrotron light sources), and so forth.
These may geographically anchor all stages
of the R&D process (16).

Let me approach my second, less conven-
tional, path forward indirectly. A 2009 Har-
ris poll (17) asked the public to name the
most prestigious occupations. The answers
(in order) were firefighter, scientist, doctor,
nurse, teacher, and military officer. What
struck me immediately when I saw this result
is that every one of these, except scientist,
is an immediate helper occupation. These
people save us from fires, prevent attacks,
teach our children, and heal us. By contrast,
the value of scientists and the benefit they
produce can be very long term. Yet the pub-
lic perceives scientists as belonging in the
same basket of high-prestige helper occupa-
tions. This tells us something. Another
poll, by Pew (18), finds that the majority of
Americans think that government invest-
ments in basic scientific research (73%) and
engineering and technology (74%) pay off
in the long run, with only small differences
between Democrats and Republicans. That
also tells us something.

When I ask nonscientists, “what is science
good for?”, only rarely do I hear the answer
that it promotes economic growth. Instead,
most answers are about science creating a
better world in ways not easily monetizable:
ensuring longer, healthier lives; protecting
the planet; feeding humanity; deterring (or
winning) international conflicts; bringing
us resiliency to engage future challenges of
an uncertain world, with climate change as
one example. People also talk about the basic
human need for discovery and understanding
of the natural world, about the almost-myste-
rious power of science at engaging the natu-
ral idealism of young people, and empower-
ing them to participate in the world in ways
that they otherwise would not.

There are numerous respected occupa-
tions that are closer to the business of busi-
ness; that is, they enable industry to create
jobs. Yet these are not at the top of the Harris
poll. My theory is that the poll results, and
also the public’s appreciation of science’s
nonmonetizable benefits, are related effects:
Both are the result of the public’s intuitive
understanding of something well known to
statisticians as heavy-tailed distributions.
When you bet on a heavy-tailed distribution
that has positive outcomes, you can not only
win, you can win big.

*What It Means for Scientific Discovery to Be Heavy-Tailed*
To the statistician, a heavy-tailed distri-
bution describes a certain kind of pattern
for how often random events with various
magnitudes occur. Heavy-tailed means that
extremely large events are only a bit rarer
than mid-sized events. In other words, the
tail of the probability distribution extends
out to the right, so as to allow events of
truly huge consequence to occur once in
a while—and more than one might think.

**“America is a country of
two wonderful, frontier, short-term
pragmatism. That benefits us
in many ways, but it may not
put the United States in the
best competitive position in
a game of international
competition that favors long
time horizons.”**
First, we can look back at Fig. 2 for historical guidance. As nations compete for economic growth opportunities, it appears that those who spend close to 3% of their GDP on R&D are the ones that compete most successfully. The United States is in that club now. We don’t want to fall out of it.

Second, we can look inward at ourselves. How patient can we be as patient investors is a matter of social and national character. It is a question of intergenerational social choice. What kind of world do we want our children and grandchildren to have? Does our planning horizon extend beyond two generations? Can we think about what kind of world we want for our great-great-great-grandchildren, who will be living a century from now? That is a very short time in the great scheme of things.

America is a country of wonderful, frontier, short-term pragmatism. That benefits us in many ways, but it may not put the United States in the best competitive position in a game of international competition that favors long time horizons. Europeans look up and see 500-year-old cathedrals. European science has developed mechanisms for obligating nation states to make investments over long periods, for example, to CERN (23), where the Higgs boson was discovered. Chinese culture sees itself as having continuity over millennia. China’s investment in R&D as a fraction of GDP is under 2%, but that number is on a trajectory of amazingly rapid and sustained increase (24).

The American public already highly values science and scientists and seems to have an intuitive grasp of heavy-tailed, not immediately monetizable, returns. Through communication with the public, we must continue to provide the evidence that may justify those beliefs—indeed, this is the mission of AAAS. But also as individuals, we must seize every opportunity to demonstrate that what we do is altruistic and idealistic and that it is also economically vital. Our message is that science is a single, unified, long-term enterprise in which basic science discoveries, and research accomplishments of applied science and engineering, are things to be admired in their own right that also, often unpredictably, lead to better jobs and better lives, new products and new industries. Both of these perspectives will be well served if the United States is able to keep itself (and help to put the rest of the world) on a Solow-inspired trajectory of technology-enabled exponential growth.
What's So Special About Science (And How Much Should We Spend on It?)
William H. Press

Science 342 (6160), 817-822.
DOI: 10.1126/science.342.6160.817

Use of this article is subject to the Terms of Service