

Responsible Citizenship in a Technological Democracy

Note 8: The “Scientific Method” and the “Engineering Method”

Theodore von Karman, a highly respected aerospace engineer from CalTech in the 40’s and 50’s, when trying to explain the difference between science and engineering, said that

Science is about *understanding* nature, understanding what *is*.
Engineering is about *creating* what *has never been*.

Understanding what *is* vs. *creating* what *has never been*. That’s a pretty good summary of the two extremes of what is actually a spectrum that many scientists and engineers move along. Pure science is analytic – it strives only to understand nature. Pure engineering is synthetic – it strives only to build things. Real science and engineering are usually a mixture of both.

Public policy often needs to be informed by our understanding of nature – just what are the causes and consequences of global climate change, for example. But it also needs to be informed about whether the effects of climate change can be *fixed*. That is, can we create technology that allows us to maintain our quality of life while also mitigating the effects of climate change? So both science and engineering knowledge are essential to good public policy.

A responsible citizen doesn’t need to know all the science or all the engineering, but does need to know what each of those can bring to the table – what their strengths and limitations are. You need to be a good “consumer” of S&E knowledge.

Understanding nature: the “scientific method”.

What we now call science used to be called “natural philosophy”, in which the method for testing whether some explanation of nature was any good was, well, philosophical. That is, one reasoned about it, debated it, and appealed to authority figures.

In the second half of the 17th century, that changed, and what we now call the “scientific method” emerged. Rather than track all the history, we’ll just consider the modern version of it – but the fundamental 17th century insight was that, instead of just reasoning about nature, we could ask nature itself about the validity of a proposed explanation. That is, we can devise “tests” or “experiments” that will have one outcome if the proposed explanation works, and another outcome if it doesn’t. Then we just let nature do its thing, and see what actually happens.

So here is the modern version of the scientific method.

1. Someone proposes a “hypothesis” that purports to explain a natural phenomenon.
2. The hypothesis is checked to ensure that it explains known phenomena at least as well as the current prevailing hypothesis. If it doesn’t, it is rejected.

3. The hypothesis is checked to see if it predicts some phenomena that have never been observed and aren't predicted by whatever the prevailing explanation is. If it doesn't, it falls into the "OK, but merely a competitor" category.
4. An experiment is devised to "ask nature" whether the new prediction is valid. If the experiment does not show the predicted phenomenon, the hypothesis is rejected. If it does show the predicted phenomenon, the new hypothesis is accepted as a better explanation of nature than whatever the prevailing one had been.
5. If the new hypothesis ousts a previous one, it precipitates a torrent of additional tests of other predictions by the new hypothesis as well as other scientists repeating previous experiments to be sure that their results can be duplicated.

Two properties are required of a hypothesis: (1) it must predict new phenomena, and (2) in the lingo of science, it must be "falsifiable" – that is, it must be possible to show that the prediction didn't happen and so the hypothesis is wrong.

In contrast to much common wisdom, scientists never know when an explanation of nature is true, they only know when one is false. Maybe that's worth repeating. Scientists never know when an explanation of nature is true, they only know when one is false. A hypothesis that correctly predicts a previously unknown phenomenon is better than the previous hypothesis, but it may fail to predict some other phenomenon, so an experiment can't prove the hypothesis is true.

On the other hand, explanations that have repeatedly predicted phenomena that were verified by experiments gain credibility. The more predictions and experiments, the greater the credibility. Hypotheses that gain great credibility are promoted to be "theories" or even "laws".

Aside on geek terminology:

In common, everyday lingo, a theory is a speculation and a law is fact. To scientists, a theory is a hypothesis with a lot of experimental evidence that it's not false. A law is a theory that has even more evidence to support it. There's no precise definition of what it takes to become a theory or a law – but the important thing for you to remember is that, while either could be overturned in a heartbeat by just one experiment showing that one prediction is not true, the accumulated evidence for a "law" is so strong that being overturned seems doubtful.

Aside on incorrect hypotheses:

Some hypotheses are *really wrong* – the Greek notion that everything was made of just four elements: earth, air, fire and water, for example. These go in the trash heap. On the other hand, some "wrong" hypotheses are useful. You don't need to know the details, but Einstein's General Theory of Relativity showed that Newton's theory of gravitation was wrong. But where the difference matters is pretty esoteric – mostly involving extreme accelerations that we don't experience here on earth. For almost all our everyday experience, Newton's explanation works just fine and is *a lot* simpler. But not for everything! The Global Positioning System (GPS), for example, has to use Einstein's General Theory to find your location correctly.

Aside on non-scientific questions:

In effect, experiments are a way of asking an impartial, and always truthful observer (nature) to resolve a debate about which of several hypotheses best explains a natural phenomenon. That's a pretty nice thing to be able to do, but one must give up something to get it. Namely, there are a lot of very important questions that can't be answered scientifically – questions about the meaning of life, whether there is a god, the correct ethical choice to make in a given circumstance, or whether to pop the question to a particular woman. These kinds of questions aren't about nature and can't be answered by the scientific method.

Creating what has never been: the “engineering method”.

For some reason, engineers have been less introspective about their discipline than scientists – so the “engineering method” is not discussed as much as the “scientific method”. But it exists, and it's important that you understand it, and both its strengths and limitations.

Engineers design solutions to human problems – solutions that satisfy society's needs or desires. So, the engineering method must start with a human need or desire and end with a solution to it. But that's a little too simplistic. Some solutions are better than others. It should end with the best solution the engineers can devise. Since engineers have been less introspective, the “engineering method” is less well defined than the “scientific method”, but here is my version of it:

1. Perform a “requirements analysis”. Determine at least three things: (1) the problem to be solved, (2) any constraints on the solution – things like size, weight, power consumption, safety and reliability considerations, etc., and (3) desirable properties of a solution – things that will differentiate better solutions from poorer ones.
2. Research the design options, including doing the iterative design through multiple “levels of abstraction” as described in Note 4 on Systems.
3. Build a prototype and test it (see note 7). What looks good in theory doesn't always work so well in practice¹.
4. Using the result of the prototype testing, refine the design to a “production version”, taking into account additional constraints of producing the solution en masse – what works for limited production doesn't always work for mass production.
5. Produce the solution, but continue testing and refining the design. As is noted below, analysis of the failures of a design is critical to improving it.

It's important that at each at each step one may have to abandon ship and drop back to a previous step. For example, when building the prototype one may find that one of the constraints identified during the requirements analysis can't be satisfied – so you need to go back and ask why that constraint was there, whether it can be relaxed or whether some other constraint can achieve the necessary effect. Once you've done that you'll have to repeat the research of design options too since some of your previous ones may not work any longer and some new options may be possible. Remember, engineering *creates what has never been* and, in the creative process, the unexpected is the norm.

¹ As they say – in theory, theory and practice are the same; in practice they aren't!

In many ways the hardest part of this is the first step – the requirements analysis. It is awfully easy to state a problem in terms of a solution and surprisingly hard to get to the real problem. For example, recall that in the list of questions in Note 3 was one that asked about the purpose of brakes in a car? Most people will answer that it is to slow or stop the car, but that’s wrong. The purpose of the brakes is to let you safely drive fast. Without them you would have to drive really slowly!

Just to expand on the point about how hard requirements analysis is – there is a wonderful book, *The Innovator’s Dilemma*, that explains how companies have a lot of trouble introducing new, “disruptive” technology. It’s because customers tend to ask for improved versions of their current technology even if a new technological approach would be better. They tend to state their problem in terms of the current solution instead of the underlying problem. *Really* rich entrepreneurs tend to be people that have the knack of cutting through to the underlying problem.

As an engineer explores the design options it’s common to encounter “trade-offs” – situations where you have to choose which of two or more constraints is more important because they can’t all be satisfied simultaneously. It may not be possible to satisfy a weight constraint and a strength constraint simultaneously, for example. NASA had a motto for a while of building spacecraft “faster, cheaper, better”. Wags at NASA changed this to “faster, cheaper, better – pick two”. What they were saying was that you can have any two of the three, but not all three. They were right. There are trade-offs!

A small aside on prototyping. Classical engineering has always involved physical prototyping – this may be changing, however. Computer simulation has gotten so good that some very complex systems, like the Boeing 777 and 787, never built a physical prototype. They built a “virtual” prototype in the computer where simulation was used to test it.

Aside

It’s interesting that the word “science” is used by scientists and engineers with two very different meanings. One is the body of accumulated hypotheses that we currently believe describes nature. The other is the process – the scientific method – by which those hypotheses were validated. When a question starts with “What does extant science tell us about ...”, it’s referring to the former – the current explanation of nature. When someone says “Jane is a scientist” they are saying that Jane adheres to the scientific method. Or, when a question starts with “Is there a scientific approach to ...” they are likewise referring to the process. Scientists and engineers seem to have no problem with this ambiguous usage, but the general public does – and that can lead to untold confusion.

By contrast, “engineering” is the process and “technology” is the result of that process, respectively. Although in some common usage “technology” has come to mean computers and digital things, in fact it is everything created by engineers: cloth, flatware, paper and books, the space shuttle, your MP3 player, potato chips, and on and on.

Policy Relevance

A common misconception of non-scientists, especially pernicious in policy contexts, is that science discovers the “truth”. Nope. It uncovers what is *not* the truth. Experiments can only show that the prediction from a hypothesis is right or wrong. If the prediction is wrong, then the hypothesis is wrong. But, if the prediction is right, that doesn’t mean that some other prediction from the hypothesis won’t be wrong. Only slowly over time, if a hypothesis withstands many experimental challenges, a consensus develops that it’s probably right. But that can be reversed in the blink of an eye by one experiment showing that one prediction is wrong. Beware the policy that requires us to know the truth about nature!

An example of this misconception can be seen in some of the current debate over climate change. Some people have argued that the cost of significantly reducing greenhouse gas emissions is so large that we should wait until science has proven that those emissions are really the culprit. Unfortunately that will never happen – it *can’t* ever happen. The best you can get is a consensus of the experts that greenhouse gasses are the best hypothesis for explaining rising temperatures. And, even then, there is no guarantee that they won’t have to reverse themselves.

As noted earlier, many policy questions aren’t about science, but rather about whether a technological solution exists to some problem, that is, about engineering. For example, the deep question about climate change can’t be whether we have a 100% guarantee of either the fact, or the cause of it. Science can never give us that. Rather it’s whether there are realistic technologies for reducing greenhouse gas emissions in the event that the current strong consensus of scientists is right that we are warming the planet in ways that threaten our wellbeing. Somehow this really important dimension of the issue doesn’t get as much public attention as much as the scientific one. Go figure.

Oddly, the limitations of the engineering method are almost the converse of those of the scientific method. Engineers know when a solution to a problem exists, but they don’t know (1) when one doesn’t exist, or (2) whether the one they have is the best possible. Being a human, creative process, the fact that no one has found a solution yet doesn’t mean that one doesn’t exist with today’s technology or that some breakthrough won’t make one possible tomorrow. Similarly, there is no way to prove that, given a solution, there isn’t one that is better – safer, lighter, cheaper, more reliable, with less impact on the environment, or whatever.

Beware of folks who come to you with the “solution” to a problem like climate change – wind, or biomass, or whatever. In many cases these are less well thought out than they seem. Use the “engineering method” outlined above as a guide to see whether they’ve done a thorough requirements analysis, for example. What were the design options considered? Did they consider the entire system of which this “solution” is a part? See if they have the data on the prototype(s). Be especially observant of failures to account for the second law: *Nothing works perfectly. Nothing!*

Another common mistake is to forget the reality of “tradeoffs” in an engineered system. Sometimes false tradeoffs are used as an excuse for a poor design, but most of the time they are a

reality that must be dealt with. Be skeptical and do a good analysis, but be prepared to acknowledge that reality. There are just times when you can't have everything you want.

Finally, the engineering method is a pretty good model for how to think about formation of public policy itself. Doing a good requirements analysis, researching the options, prototyping, etc., are all good mental discipline for the formation of policy. In particular, just as with technological solutions, the hardest part of policy formation is often getting the requirements analysis right – too often people start with the “solution” without having thought deeply about what the problem is!