The Effect of Peer Review on Progress: Looking Back on 50 Years in Science

Thomas Gold, Sc.D.

The pace of scientific work continues to accelerate, but the question is whether the pace of *discovery* will continue to accelerate. If we are driving in the wrong direction–in the direction where no new ideas can be accepted–then even if scientific work goes on, progress will be stifled. This is not to suggest that we are in quite such a disastrous position, but based on my personal experiences during more than 50 years of work in various branches of science, I fear that all is not well.

What Motivates a Scientist?

A scientist, in my naïve definition, is a person who will judge a matter purely by its scientific merits: the evidence as it currently stands. His judgment will be unaffected by the historical evaluation of the subject, or by his perception of how his conclusions will received by his peers or how they will influence his standing, his financial position, or his promotion.

I may have reduced the number of those whom you think of as scientists very considerably, even to a null class. But we have to be realistic and realize that people have certain motivations. The motivation of curiosity is an important one, but I doubt that there are many scientists to whom this motivation would suffice to go through a lifetime of hard struggle to uncover new truths, in the absence of other reasons that would drive them along that same path. If there were no question of acknowledgment, a reasonably comfortable existence, and so on, I doubt that many people would choose a life of science.

These other motivations may of course cloud one's scientific judgment. And we must ask about communal judgment-clouding factors. What is the effect of the sociological setting? Is our present-day organization of scientific work favorable or unfavorable in this respect? Are things getting worse, or are they getting better?

One Barrier to New Ideas in Science

New ideas in science are not always right just because they are *new.* Nor are the old ideas always wrong just because they are old. If we look over the history of science, there are very long periods when the uncritical acceptance of the established ideas was a real hindrance to the pursuit of the new.

In preventing the acceptance of scientifically valid new ideas, one obvious factor that has always been with us is the unwillingness to *learn* new things. I can give you an example from my own experience.

When I was still very young, just after the Second World War, I had worked on the theory of hearing: how the inner ear works. As I had just come from wartime radar, I was full of signal processing methods and sophistication and receiver techniques and all that, and there I found myself discussing the physiology of hearing in those terms. I thought it was very appropriate because it is a very fine scientific instrument that we were discussing, the inner ear. But I had to address myself to an audience of otologists, the only people who were doing any kind of research in this field. The mismatch was obvious; it was completely hopeless. There was no common language, and of course the medical profession just would not learn what it would take to understand the subject. On the other hand, they certainly made their judgments about the matter. This reaction essentially forced me out of the field.

The theory of hearing that I proposed then involved an active–not a passive–receiver, one in which positive feedback, not just passive detection is involved. We now have very clear evidence that indeed an *active* receiver is at work.

But reluctance to exert the effort to learn something new is by no means the most serious problem.

The Herd Instinct

In tribal society, what I call the "herd instinct" presumably has some sociologic value. In science, however, we generally want diversity–many different avenues need to be pursued. When people pursue the same avenue all together, they tend to shut out the other approaches–including what may be the right one.

A flock of starlings may be able to change course all at once, but this generally doesn't happen when one member of a herd decides to head in a different direction. If a scientist adopts another viewpoint, there is no way to be sure that others will follow. He may be left alone outside the herd. Moreover, he will be challenged to justify why he has departed. The others will never be asked why they stayed. The sheep in the interior of the fold are well protected from the bite of the sheep dog.

I sometimes wonder whether the much encouraged and acclaimed interaction among western astronomers leads to a form of mental herd behavior that, if it does not actually put a clamp upon free thinking, insidiously applies the pressure to follow the fashion. This made the writings of colleagues in the former Soviet Union, who developed ideas in comparative isolation, all the more valuable.

Indeed, I have wondered whether one should in fact pursue subjects with a big wall between two groups that are working in the same field, so that they absolutely cannot communicate, and see a few years later whether they come even approximately to the same conclusion. It would then give some perspective of how much the herd behavior may have been hurting. But we don't have that. Even with our Soviet colleagues, unfortunately, we had too much contact to have a display of real independence, to see where it would have led.

Peer Review Demands Conformity

The herd instinct in individuals is augmented by the support structure of the scientific enterprise–not only the financial support but also the journals, the judgment of referees, the invitations to conferences, and acknowledgments of every kind. The scientist who lacks the approbation of peers will suffer consequences.

It is important to recognize how strong the interaction of the support of science and herd behavior really is. Suppose that you have a subject in which there is no clear-cut decision to be made between a variety of opinions and therefore no clear-cut indication of the best direction to favor in funding or publications. No doubt opinions would need a multidimensional space to be presented, but I will at the moment just represent them in a one-dimensional situation.

Suppose you have some curve between the extreme of *this* opinion and the extreme of *that* opinion. You have some indefinite, statistically quite insignificant distribution of opinions. Now in that situation, suppose that the refereeing procedure has to decide where to put money in research, which papers to publish, and so on. What would happen?

If a few more people believe in one position than in any other, speakers at the next conference will be selected from this group, and it will receive more funds. Remember that the referees themselves are on the curve.

A year later, what will have happened? Some of the people in the disfavored region will have been filtered out, and more will be in the favored region. Each round of decision making has the consequence of essentially taking the initial curve and multiplying it by itself.

Now we understand the mathematical consequence of taking a shallow curve and multiplying it by itself a large number of times. What happens? In the mathematical limit it becomes a delta function at the value of the initial peak.*

If you apply the process long enough to a probability distribution function representing opinions, you will have created the appearance of unanimity. It will look as though a problem has been solved because all agree.

Although the peer review system is widely regarded as the only fair way to distribute grants, it inevitably manifests this statistical effect. The more reviews required for a proposal, the more certain is the effect. If there were noise in the situation, it would be much better. There used to be many different agencies in the United States, and there was perhaps an odd-ball over here who gave out some money for one agency, and a funny fellow over there for another. This was a noisy situation, and it was not driving quite as hard towards unanimity. But now we have it all streamlined and know exactly to whom we have to go for a particular subject and, of course, the entire herd now has to move in the same direction.

Why is it thought that the peer review system would work for science? How about trying to make a peer review system work for other forms of endeavor? Suppose we had a national foundation for the arts, and every painter had to apply to it to get his canvas and his brushes and his paints. I can imagine some of the consequences, but better than that, we can look them up in historical examples.

Eduard Manet wrote to his colleague Claude Monet, of Renoir: "He has no talent at all, that boy. Tell him to give up painting." Degas regarded Toulouse-Lautrec as "merely a painter of a period of no consequence."

* The Dirac delta function is an odd mathematical beast applicable to things such as the charge density of a point particle. It may be quantitatively visualized as a limit of a process that starts with a square barrier and reduces its thickness while increasing its height so that the area under the barrier remains 1. The value of the function is zero almost everywhere but infinite at the point at which the barriers were centered.–*Ed.*

Peer review in music would be no different. "An orgy of vulgar noises" was the verdict on Beethoven's Fifth Symphony by Mr. Spore, a German violinist and composer. Of Brahms, Tchaikovsky said, "What a giftless bastard. It annoys me that this jumping, inflated mediocrity is hailed as a genius."

The book *The Experts Speak: the Definitive Compendium of Authoritative Misinformation* by Christopher Cerf and Victor Navansky also gives many additional examples.

Thus we see that the herd instinct is a pervasive tendency in the human makeup. While this is a severe handicap for science, we have arranged a system that strengthens this instinct instead of combating it as best we can. It is virtually impossible to depart from the herd and continue to have support, a chance of publication, and the other advantages that one requires to work in a field.

Every journal has to send each article out to a number of people to review, and most of the people are with the herd. Usually with just one-third of the reviewers very negative, the paper does not get published. Likewise, there is no free speech at conferences. With rare exceptions, those with a divergent opinion cannot raise funds to run their own conferences. In universities, even if the Dean is willing to promote somebody who is outside the pack to tenure, he cannot do so because he must send letters to the leading persons in the field to get permission.

The herd problem is especially bad in the planetary sciences. The National Aeronautics and Space Administration (NASA) made the grave mistake not only of working with a peer review system, but one in which some of the peers (in fact very influential ones) were the in-house people doing the same work. This established a community of planetary scientists that was completely selected by the leading members of the herd, and was very firmly controlled, and after quite a short time, the slightest departure from the herd was absolutely cut down. For all the money that has been spent, the planetary program will one day be seen to have been extraordinarily poor. The pictures are fine as are some of the facts that have been obtained from the planetary exploration with spacecraft, but little else will stand.

Breaking a Consensus

Once a herd has been established in a subject, it can only be broken by the most brutal confrontation with opposing evidence. There is no gentle way that I have ever seen in the history of science that has been successful.

In many subjects such clear evidence is very hard to come by. In the complex subjects–I always think of the earth sciences in this respect–there are always different ways of interpreting any one fact because so many complicated things have taken place. All the money spent obtaining evidence may be wasted, or worse, may actually serve to cement further the bad situation. So it is very likely that money is often spent in science in a way that is absolutely detrimental to that science.

In early 1960s or late 1950s, any application to investigate the possibility that continents are moving around a little would have been instantly rejected by referees. That was the notion of a crackpot. Six years later you could not get a paper published that doubted continental drift. The herd had swung around–but it was still a firm and arrogant herd.

Shortly after the discovery of pulsars I wished to present an interpretation of what pulsars were at the first pulsar conference–namely that they were rotating neutron stars. The chief organizer of this conference said, "Tommy, if I allow for that crazy an interpretation, there is no limit to what I would have to allow."

I was not allowed 5 minutes of floor time, although I in fact spoke from the floor. A few months later, this same organizer started a paper with the sentence, "It is now generally considered that pulsars are rotating neutron stars."

The discovery of contrary evidence may, however, not be sufficient to turn the herd: there is also "shoehorn science." When in a subject a general attitude or a viewpoint has become established, then it is very easy to obtain funds for a proposal that states: "I will demonstrate how this fact and that fact, which are apparently difficult to see in the accepted framework, can be made to fit into that framework." And by the time that much work of the shoehorn kind has been diligently done to force the facts into the preordained pattern, it then looks to many people as if it were firmly established. A large superstructure may be built on what may actually be no foundation. If I may invent a "Confucius say" proverb, "Never judge strength of foundation by size of building."

Petroleum geology is a prime example, from which I learned that if one dares to look at the foundation, one is immediately a scoundrel. People became absolutely wild and shook their fists at me when I proposed in my talks that there was some uncertainty about the origin of petroleum. And how could one obtain funding to explore that prospect?

A colleague and I were able to obtain unsigned referees' reports on an application we presented to the U.S. Department of Energy to fund a study investigating the chemistry of hydrocarbons at high pressures and high temperatures in the conditions in which they might be at some depth in the earth. One wrote, "This proposal *must* be funded. In science every research project is a risk, but here the risk is negligible because even if the hypothesis is not correct, this research proposal will contribute strongly to fundamental science in petroleum engineering, the thermodynamics of fluids, and geochemistry. If the hypothesis is correct, the Department of Energy will have hit the jackpot beyond its wildest imagination." And he continued with the detailed questionnaire with top marks in every part: the competence of the proposer, the institution, the test, the facilities, and all that.

There was a second referee who also gave it top marks for all the questions. But in response to the last question, "Should this proposal be funded?", he wrote, "No." And in answer to the question "Why not?" there was a single word: "Misguided." Because of that word and similar words by two or three other referees, the project was not funded.

Because of many such experiences over the years, both with the National Science Foundation (NSF) and the Department of Energy, I concluded that it was absolutely hopeless to get any money in contravention of the opinions that are so firmly established in the petroleum business as to how oil and gas came to be where they are. The fact that we find oil and gas on other planetary bodies, obviously not produced by biological factors specific to earth, was usually ignored completely, although one fellow actually got a paper published asserting that there *must* be life on Jupiter because hydrocarbons have been found there. Finally, I was able to get money from the gas industry to do research on this subject.

The Herd Effect in Industry

In a field such as petroleum that involves a large number of people because of the economic applications, the problem of the herd instinct is aggravated. First, there are many mediocre people in the field, who overpower it by sheer numbers. Then there are powerful disincentives to admitting error. The petroleum geologist who has been advising Exxon how to spend hundreds of millions of dollars for 30 years is unlikely to go to his bosses and say, "I am sorry, Sir, but I have been wrong all those years. We have been finding petroleum, but if we had searched for it in another way, we would have found 10 times as much."

As Tolstoy wrote:

> *I know that most men, including those at ease with problems of the greatest complexity, can seldom accept even the simplest and most obvious truth, if it be such as would oblige them to admit the falsity of conclusions which they have delighted in explaining to colleagues, which they have proudly taught to others, and which they have woven, thread by thread, into the fabric of their lives.*

Is There Another Way?

The peer review system may make people reasonably happy, and may be regarded as fair, but it doesn't work. We must ask whether a better system is possible.

The best method I can think of is similar to what Arthur Katrowitz proposed at least for major decisions: the "science court." When much is at stake, and a subject has been driven into an alley, one must set up a science court in which the different viewpoints would be argued by the protagonists of each one, with carefully prepared work. The different viewpoints could be judged, not by others working in that same field–as that would merely take you back to the herd–but by a group of very knowledgeable and very competent scientists distributed over other fields, but with enough general competence to be able to listen and understand the detailed arguments of the field in question. I would be much happier to have subjects surveyed every now and again by a jury of that kind. It has to be a scientific jury because it would have to understand detailed scientific arguments, but members do not have to be–and should not be–from the field in which the decision is to be made.

I propose that in every field the NSF should set up such a science court to hear all the different opinions on a reasonably regular basis. While this could not be done for every application, it could be done sufficiently often for major decisions to break, or at least spoil somewhat, the herd system.

Without some mechanism of review by independent persons outside the herd, a very large proportion of science funding will remain firmly in the wrong hands.

Science is the search for truth. When money and fame are on the line, the truth suffers. We need a method to correct for the flaws in human nature, instead of one that magnifies them; a method to correct errors promptly, instead of propagating them. We need a culture that encourages the virtue of humility, found most prominently in the greatest scientists, and the openness to consider the possibility that we just might be wrong.

Thomas Gold, Sc.D., is Professor Emeritus of Astronomy at Cornell University. He was the founder and for 20 years director of the Cornell Center for Radiophysics and Space Research. E-mail: tg21@cornell.edu.

***Editor's note**: This article is an edited and updated version of a speech, an unedited transcript of which appeared in the* Journal of Scientific Exploration *1989;3(2):103-112, prepared and published with the author's permission.*