Perspectives Editor’s Note
Dialogue that openly addresses current issues and questions what has been considered as dogma is critical for scientific advancement. Here, a successful entrepreneur, author, and Professor of Practice at the University of Bath School of Management provides valuable lessons on how to approach scientific research with the goal of having lasting impact through examining what we do to find true and lasting value in our work.

EXPECT THE UNEXPECTED

Margaret Heffernan
Professor of Practice
University of Bath UK

Models and Tools

In the early 1950s, an unusual scientist named Alice Stewart was looking for a research project. The burning issues of the day – lung cancer, cardiovascular disease and polio – were crowded. This left leukaemia. One anomaly in particular caught Stewart’s eye. Disease is usually correlated with poverty, but in this case, the children dying from leukaemia came from countries with better medical care and lower overall death rates. That didn’t make sense, so it was intriguing.

Stewart decided to interview the mothers of leukaemia victims to see if she could find what might account for this pattern. She didn’t really know what she was looking for and, with only a tiny grant, she had just one chance to survey the parents of 1,000 children who’d died of cancer and 1,000 who had not. So she had to ask about everything she thought might have an effect, including exposure to infection, inoculation, cats, dogs, chickens, fast-food consumption, high-sugar soft drinks, candies, and X-rays [1]. By a rate of three to one, exposure to X-rays was associated with cancer. The result had a statistical clarity most scientists can only dream of.

When her article entitled ‘Preliminary Communication: Malignant Diseases in Childhood and Diagnostic Irradiation in utero’ that attempted to trace all children in England and Wales appeared in the Lancet in 1956, it caused a stir [2]. The Nobel Prize was mentioned and Alice was asked to conduct a similar survey in Scotland. In three years, Stewart and her small team had traced 80% of all childhood cancer deaths in the UK between 1953 and 1955. Their report in the British Medical Journal in 1958 concluded that a foetus exposed to an X-ray was twice as likely to develop cancer within the next 10 years as a foetus that had not been exposed. They estimated that a child a week was dying from the practice. Yet the British and American medical establishments continued to X-ray pregnant women for the next 25 years.

Why did a scientifically-literate audience ignore the data? Many assume the answer to be gender discrimination but that doesn’t hold; Brian MacMahon’s subsequent, larger study at the Harvard School of Public Health confirmed Stewart’s findings [3]. What prevailed was threshold theory: the dominant model of disease at the time, this maintained
that everything was safe up to a point, a threshold, only after which it became dangerous. Stewart and MacMahon argued that there was no safe level of radiation exposure for a foetus. But asked to choose between their model and the data, the medical establishment stuck with their model.

The dazzle of new technology was also implicated. Ever since their discovery in 1895, X-rays had developed an aura of mastery and mystique. They were used as an exquisite and expensive form of portraiture in the 1890s and to find a ring that had been mistakenly baked into a cake. Shoe shops boasted of X-ray machines that ensured a perfect fit while doctors used the machines to find bullets, identify broken bones and tumours. Many physicians invested in them as a premium tool in their clinics. The big tech of their time, X-rays were too helpful and too appealing to be questioned.

Today, nobody X-rays pregnant women except in cases of dire emergency. So it’s easy with the benefit of hindsight to see that, in clinging to their model, the doctors made a poor choice. But we all use models of how the world works and we cling to them. They make sense of our knowledge, help us to weigh new information and to discriminate signals from noise. They frequently reveal areas of vast potential which in turn generate new disciplines and discoveries: what to work on next, where the next good question lies? In that respect, they demarcate fruitful from arid territory. The longer we spend inside our models, the more attached we become to them and the more useful they become to us. But in so doing, they develop in us a strong bias for perspectives and inquiries that confirm the model while repelling, marginalizing, and trivializing disconfirming data.

Given enough institutional enthusiasm, models can become fads, heavily invested in by research bodies and commercial companies. DDT (Dichlorodiphenyltrichloroethane) was enthusiastically developed as a new model of warfare, then applied to agriculture and then to domestic gardening. Only later was it found to be an environmental disaster. The development of TIBA (2,3,5-triiodobenzoic acid), which accelerates flowering in soybeans, later became the basis for Agent Orange. The creation of the World Wide Web was first as a scientific tool, then a publishing tool, then a data aggregation tool used to profile and surveil individuals and groups. The decoding of genomes facilitated genetic testing, and then CRISPR which enables the genetic engineering of human beings. The rapid discovery of knowledge and application is frequently accelerated by enabling technologies that, together with models, create the scientific equivalent of stock market bubbles, where wild excitement disables both critical appraisal and all sense of personal responsibility. While these heated moments often generate important breakthroughs, they also carry with them the danger of irrational exuberance: the failure to see or to take responsibility for unexpected consequences.

Clinging to threshold theory, combined with wild enthusiasm for X-rays, cost thousands of children their lives. Not because scientists didn’t care about children, but because the allure of the model and of the technology went unquestioned for so long. That there are so many instances in which this happens, where the harm is seen only in retrospect, forces the question: how might scientists and their research teams develop proactive habits of mind and practice that mitigate mindless enthusiasm and make space for responsible reflection and challenge?
“Doubt is uncomfortable but certainty is absurd” (Voltaire)

Developing doubt is an essential ingredient in all innovation. Before important decisions, many teams and organizations adopt the habit of pre-mortems: if this could go wrong, how would it go wrong? What might the consequences be? Challenging underlying assumptions can be helpful too. In research teams where everyone adopts the same mental model, assumptions need to be challenged, questioned and sometimes updated. Doing so is far rarer than it should be; everyone is eager to crack on. Simple questions – if we had more time, what would we want to know? If we had less time, would we feel confident enough to proceed? – can unpack discomfort and useful doubt. But such interrogation rarely comes easily to most people.

Conducting this kind of negative exploration is difficult but essential. Research into organizational silence shows that in many industrial accidents or organizational failures, catastrophe strikes not through an absence of knowledge, but the lack of discussion. The two prime reasons for this silence are fear of conflict (‘I may start an argument I don’t know how to conduct and I might lose’) and a sense of futility (‘I could raise the issue but it won’t make any difference.’) Such fears account for up to 85% of concerns that remain unvoiced [4]. That means that most of the knowledge, insight and questioning inside the minds of co-workers is not voiced. So silence should never be taken for assent; rather it signals the need to find safe ways in which to open up debate.

When there is a great deal at stake, these kinds of debates are often formalized, with each participant assigned a specific approach. One might consider the impact of more or less time, more or less funding, more or fewer collaborators. How much of the decision needs to be taken at once? What are the unknowns, how dangerous are they and can they be reduced? Under what circumstances would this look like a very bad choice?

For critical questioning to become endemic in research requires going deeper. Research progresses through teams because collectively groups of people make better decisions and see better solutions when working together than working alone. But what distinguishes great teams from weaker ones? Experiments led by Thomas Malone at MIT [5] found that IQ was not a determining factor; getting the smartest guys in the room together didn’t guarantee success.

Instead, collective intelligence depended on three factors. First, the collective score on the *Reading the Mind in the Eye* test. This is roughly considered a test of empathy: how connected am I to you, how deeply do I think about you as you? Second, equality of contribution. Teams in which there are no passengers but also no dominant voices tend to solve more problems more easily. Third, gender diversity. Teams with more women in them also tended to perform better. Whether this is because women bring different perspectives or because they tend to score more highly on the *Reading the Mind in the Eye* test (and so increase the empathy quotient), the researchers aren’t sure. But what’s so striking about these conclusions is the implication that what matters most is what happens *between* people.
Sociologists call this social capital: norms of generosity, reciprocity and trust. Where individuals feel safe, surrounded by generous colleagues who are trusted, trustworthy and who help each other, it is far easier to ask challenging, critical questions in the understanding that they will be interpreted as helpful rather than destructive or competitive. Challenge is not an attack but a gift, offered to forestall error and reinforce excellence and accelerate innovation.

The development of social capital requires time, for people to know each other not merely as colleagues but as people. It is well established that teams get better over time but a more than purely professional understanding of each other’s work matters too [6]. This is one reason why the biologist Uri Alon insists that, in his weekly two-hour lab meeting, half an hour be devoted to non-science subjects [7]. Celebrating birthdays or anniversaries, talking sport or arts or politics develops a sense of colleagues as whole people, not just co-workers or rivals. This, he told me, is what keeps people going when the work gets difficult. That solidarity facilitates constructive conflict, argument, debate of a kind that can illuminate hazards and reduce groupthink. (In this context, it’s important to remember that the salient insight of Irving Janis’s groupthink thesis is that likeminded people are more inclined to take greater risks [8].

Much creative work (in which I include scientific research) requires conflict. The Wright brothers described themselves famously as ‘great scrappers’. Their arguments began as vague ideas which were refined through debate, alternatives, fault-finding and solution-seeking. They both understood that what might have looked to others like aggression was in fact a manifestation of their shared desire to achieve breakthroughs and that that was better achieved with questions than compliance.

Questions, challenge, and debate achieve several ends. They may help to reframe a topic and thus reveal otherwise obscured obstacles and risks. They may also simply impose a pause, a halt to giddy momentum. Carlo Rovelli writes revealing about power of hesitation noting how both Darwin and Einstein are unsure of themselves as their great theories begin to emerge [9]. They do not jump in. This is the opposite of the eureka moment; instead it’s a moment of slowing down, taking time, making sure. It has much in common with the neuroscientist Mariano Sigman’s observation that competitive chess players make a note of their proposed moves before making the move itself [10]. This costs precious time, but it provides them with a moment to reflect before they act. This cooling off moment can be precious; that it is written into significant contracts testifies to the shallowness of haste.

To mitigate the costs of organizational silence, I now believe that, in all walks of life, it is vital to teach young professionals how to raise their questions and concerns in ways that won’t leave them ostracized or punished. This is harder to do than it sounds. The current generation of students is generally unaccustomed to perspective-taking, a key ingredient in constructive conflict. If I don’t know how you think and what you value, I’m unlikely to be able to frame a question in a way that wins respectful attention. Craving certainty, young students generally avoid risk and confrontation. To learn how to articulate their doubts and concerns requires that they trust their supervisors to listen with open and enquiring minds. That trust depends critically on their having seen this in action for themselves frequently.
“Time will say nothing but I told you so”. (Auden)

In the Automobile Museum in Turin, one of the very earliest cars in the world is electric. Looking at it, it’s impossible not to think: if only that model had been the foundation of this new industry, our climate might be in a wholly different state today. But the adopters of the internal combustion engine could not imagine the consequences of their preference.

The history of science is littered with regrets. Nobel and dynamite. DDT. Agent Orange. The internet and surveillance capitalism. Nuclear physics and the atom bomb. The X-raying of thousands of pregnant women. Some consequences could have been foreseen and stopped, but mental models, hype, fear, and ambiguity obscured them. These I call examples of ‘wilful blindness’, where different needs and impulses got in the way of responsible intentions. Wilful blindness is originally a legal concept, in which if there is information that you could or should know, but manage not to know, the law holds that you are culpable for negative consequences because you had the opportunity for knowledge but shirked it. As such, this is a tough standard, but a pragmatic and useful one.

But there are other moments in which the long term consequence of a research project may be unknowable. At the beginning of the eighteenth century, was it possible for Abraham Darby to see coke smelting lead to the development of a climate crisis? It seems implausible [11]. The very earliest pioneers of nuclear physics could not have foreseen the atom bomb yet every physicist in the field had to grapple with having contributed to the destruction of Hiroshima and Nagasaki. That there is no easy test for the consequences of an invention does not grant an automatic licence to abjure responsibility. The moral impetus therefore must be to seek to explore and understand consequence as best one can. To look under rocks. To tolerate the scepticism and doubt of peers and juniors, the younger generations who will inherit the benefits and/or damage of emerging work.

Scientists often sidestep this question by arguing that science is not fundamentally driven by the genius of individual scientists. What is discovered will be discovered, if not by one researcher than another. Leibnitz or Newton, calculus would happen regardless of personality. Kurt Diebner or Robert Oppenheimer, Einstein or Heisenberg, the bomb would have emerged anyway. This positions science as some form of manifest destiny for which no one is accountable. It also treats scientists, in effect, as children or slaves, witless in carrying out work they don’t understand. This might come close to how the work sometimes feels, but it earns neither plaudits nor respect from a public that funds the work or the populations which must live with its results.

A better approach can be seen in the work of the late Mary Warnock who chaired the Committee of Inquiry into Human Fertilisation and Embryology in Britain from 1982 – 4. At the time, the combination of scientific discovery and emerging technologies made it possible to start thinking about ways of treating infertility. But what would that mean for women, for families, and for society? How much change might this propel and what were the limits of public tolerance and understanding of it? Debate raged among
scientific groups and market forces started to heat up at the same time as significant opposition began to demand prohibition. Numerous unnamed fears ran amok in the public imagination.

Warnock’s great accomplishment was to convene a group of scientists, ethicists, lawyers, and doctors to explore what could be done to support the promise of the science without leaving the public angry, confused, or uninformed. Being a philosopher was her greatest advantage, Warnock said, because philosophers have no subject of their own [12]. Their craft is exploration. And her work remains admired today because it achieved two difficult things: it put definitions and boundaries around what was possible which people – citizens and scientists – could accept and could understand. The infertile were not denied the benefit of science and neither were scientists treated like pets who were left running wild because they could not be expected to understand responsibility.

In the intervening years, the development of inclusive deliberation has refuted the belief that the public can’t understand complex issues or argument. In the deliberative experiments of James Fishkin [13] and the hundreds of experiments in deliberative assemblies around the world, the capacity to articulate deeply complex ideas, technologies, models, and trade-offs has demonstrated that citizens, when treated with respect, can and do make intelligent, sound decisions, even in areas in which they are not expert. These processes are not debating forums, nor are they full of polarizing rhetoric from vested interests. Instead, meticulously managed, they ensure clarity, understanding, and constructive conflict. They are designed to produce what Warnock created: understanding and acceptance. But they go one step further than she did: they bestow on new science a degree of legitimacy which scientists cannot claim for themselves.

In 1997, 41 years after Alice Stewart’s initial research was published, Richard Doll rather quietly retired threshold theory [14]. Nobody at the time thought to estimate its value, or its cost. Since then, the world of scientific research has changed greatly. The world expects more transparency from its scientists and a greater capacity to communicate and collaborate. Who knows how many future generations may inherit the rewards of a more open and creative dialogue?

References


