JR-BRU 5509-03

How quantitative evidence can be used irresponsibly by responsible scientists

Jerome Ravetz

Presentation to the workshop,

Brussels, 9-10 June 2015.

My job as the first workshop speaker is to reflect on the seminal report by our colleagues Giampietro and Saltelli, that is the basis for our discussions. As expected, they have described the problem brilliantly, and they have offered a solution that I believe can provide the elements of reform. As they show, there is much that is wrong, sometimes seriously wrong, in the provision of scientific advice for policy. I hope to consider their analysis, and if possible to tease out some insights about the causes of the situation. That could give us some idea of its seriousness, and then, following that, of the requirements for any reform to be successful.

First, I want to remind us of the pioneering role of the Joint Research Centre in supporting critical reflection on the problems of science in the policy process. Reflection does not come naturally to any institution, least of all one that has a measure of responsibility for what happens in the world of policy formation and implementation. And yet the JRC has provided a sheltered environment for the constructive criticism that is very necessary for the institution, and indeed for policy-related science in general, to maintain its relevance and health in these turbulent times. From the time of his arrival, our truly distinguished colleague Silvio Funtowicz has promoted a methodological critique of scientific argument. It has been my privilege to collaborate with him, first on the NUSAP notational system, and then on the theory of Post-Normal Science. It is hard to imagine any other institution where he would have been given such support, both moral and material. And without such support, it is most likely that our work would have remained out on the far margins of attention. As it is, the NUSAP approach has been adopted by the Netherlands Environment Assessment Agency for its reports, in its 'Guidance'; and Post-Normal Science is now recognised as relevant by a leading Chief Science Advisor, Sir Peter Gluckman in New Zealand. The tradition has continued up to the present, as with the report by Angela Pereira and Andrea Saltelli on the problems of the JRC, and indeed with this present workshop. This is really a unique achievement for a science-policy institution, and we all hope that the atmosphere of support will continue into the future. The need for such critical reflection will not go away, and as yet there are only a few centres, I think of Bergen and Phoenix, where our sort of work is fostered. I hope that the success of this workshop will be strong evidence that the endeavour is worthwhile.

I need not spend much time reviewing the workshop document. We all know that there are very severe problems with quality of scientific materials, both for policy and even within research science itself. The ideology of 'speaking truth to power' is now obsolete, even antique. But where do we go from here? We certainly need the threefold approach to quality, as the report recommends.

But that will certainly involve an 'extended peer community'. How will the mainstream scientific enterprise respond to such a suggestion? However it is implemented, it would involve a ceding of power and legitimacy to some external agencies. Such changes are unwelcome, wherever they occur. Those who are expected to share power and legitimacy must come to see that this is certainly a least-worst solution, and hopefully one that will improve their own quality of work and life. For that, we will need a dialogue, and that dialogue must involve a deeper critique of the dominant practice, so that it is universally seen as unsustainable, both practically and morally. You may find it inappropriate for me to take lessons from Northern Ireland and South Africa for the reform of science, but the principles of nonviolent change are universal.

My contribution today is to sketch the elements of a possible version of that deeper critique. It must be along non-violent lines, not accusing anyone of deliberate malpractice. Rather, by showing that even well-intentioned and responsible scientists find themselves behaving otherwise than they would wish, we can raise the level of analysis to dysfunctions within the scientific system itself, even to the level of its defining, but usually implicit, ideology. To a great extent my work has already been done for me, by the essays in that visionary collection, *Science, Philosophy and Sustainability: The End of the Cartesian Dream*, edited by our very own Ispra colleagues, Angela Pereira and Silvio Funtowicz. My thesis, put simply, is that the persistence of that Cartesian dream, of achieving knowledge and power in a complex world by means of a drastically simplified picture of reality, is now grossly counterproductive.

Our colleagues at the workshop will be showing us important examples of widespread incompetence and indeed malpractice in the use of mathematics for policy. All too many insignificant digits are bandied about in that area. There has already been one nearly catastrophic incident resulting from the abuse of mathematics, the 'credit crunch' of 2008-9; and we do not know when and where the next one will hit. My task here is to explain how responsible scientists can systematically work with inappropriate mathematical tools and spurious quantitative material, never seeming to notice that something is wrong. There are all sorts of explanations for this unwitting corruption, which have been explored by colleagues here. Just now I want to consider philosophy. For at the heart of the modern scientific programme stemming from Galileo and Descartes, is the faith that with quantitative inputs and mathematical techniques, nothing can go wrong. They were quite clear about this, and their commitment has been echoed by mathematical scientists down through the generations up to now. The existence of this faith is most easily seen in the case of the mathematical social sciences and technologies, particularly economics, the mathematical decision sciences, and even the sciences of complex systems that rely on quantitative indicators for their data. The Cartesian faith also manifests in a less sophisticated, but totally pervasive assumption, that numbers, that is quantities expressed in digital form, are nuggets of truth. By their form alone, they guarantee accuracy through their precision. Stated baldly here, that might seem extremely naïve; but a glance at the vast scientific literature in would-be quantitative fields of complex realities shows that that faith is active and indeed hegemonic.

To interpret my examples as evidence for a Cartesian fantasy, I invoke a principle about low-quality workmanship. This is, that when a product is 'peculiarly bad', such that no competent practitioner could have produced it and yet one did, then the fault lies not with the practitioner but with the system in which they worked. My first example is of a sort of blunder that is common to the point of being universal, that is excess precision that can be called hyper- or even pseudo-precision. It is not difficult to find cases that are even quite ludicrous. One such occurs in a table of entries relevant to climate change. These are national annual emissions of CO2, either direct or including the estimate effects of land use change. A moment's reflection establishes that such quantities are highly inferential. Their uncertainties must be great, particularly for less developed countries with less developed monitoring facilities. So I experienced first surprise, then amusement, and then bewilderment when I saw that the table of entries uniformly listed all quantities to the first decimal place. That was a tenth of the hundred-million-tons represented by the entries. In the case of large countries like China, that would be a part in several billions. Such precision is meaningful only in some exceptional laboratory contexts. In the case of less developed countries, as the Democratic Republic of Congo, the precision was only one part in some hundreds of thousands, less ludicrous on the face of it, but even more so as science. My first thought was, what sort of incompetents must they be, to produce such misleading nonsense? Then the case got really interesting, as I read in the methodological appendix that the authors are fully aware of the uncertainties, estimating them at ±50% in some cases. Given that level of awareness, how on earth could responsible scientists allow themselves to be involved in the production of such low-quality, misleading material? One can only suppose that the specification of the contract for the data included such pseudo-precision, to the first decimal point, as mandatory. The scientists then had the choice between submitting good data which they knew would be abused, or leaving the job to others who would probably deliver total rubbish! But then, why would anyone want to publish such anti-scientific material? Partly because it made a tidy table that anyone could use without being intimidated by special conventions and explanatory footnotes. And also, I speculate, because in the last resort they held to the Cartesian faith that numbers cannot lie, but necessarily deliver truth, the whole truth, and nothing but the truth.

Let me pass to my second example, which is more sophisticated and deals with the mathematics of uncertainty. This is a formula for calculating the probability of the joint failure of two 'products' or securities. It achieved fame, being featured in *Wired* magazine as 'The Formula That Killed Wall Street'. It played a significant role in the destruction of speculative value in the recent crash. In the midst of the orgy of fantasy, greed and computer technology that characterised that episode, a single formula could not have been the sole agent of collapse. But it was significant. And most interestingly for our present purposes, its creator, David X. Li, is a brilliant mathematician who was appalled at the abuse of his invention by incompetent practitioners. But its very name gives the game away: 'The Gaussian copula function'. Now, 'function' is not problematic; and 'copula' is just a fancy word for 'connection'. What about 'Gaussian'? Ah, there is the secret to the success of the formula, for it enables the joint

probability of a failure to be calculated. All we need to do is to input the parameters of the Gaussian distribution of uncertainty for each of the two 'products', and the formula yields the parameters of the Gaussian probability of their joint failure. Did I say, Gaussian probability? That refers to the great mathematician Gauss, who created the distribution variously known by his name, or as 'normal'. It has two great advantages for practice. One is that it applies well enough to a great variety of number-sets in practice, particularly the averages of random deviations from a mean value of empirically derived data. And it is also very elegant in its mathematical derivations and applications. So what would be more natural than for the mathematician Dr. Li to assume that the uncertainties in his products can be characterised as Gaussian?

That assumption turned out to be disastrously wrong, hence the title. It was also ludicrously wrong, rather like my decimal place for the CO2 data. And in that way it is very interestingly wrong. For the 'products' to which the formula was applied are those financial creations which were frequently so contrived and arcane, that the investment people gave up on understanding them, and just trusted the 'quants'. Imagine combining 'derivatives' (bets on something happening, or not) with insurance policies on the fate of the investments in those derivatives, and iterating to insanity. By the end, it is impossible to put a reality-based value on such things; it is more impossible to estimate the uncertainties around that imposed value; and to assume that those uncertainties are a well-behaved probability distribution with quantified parameters – lies beyond the furthest reaches of reason. And yet, the whole financial community put its faith in such formulas. Warnings, from responsible critics including Li himself, were brushed aside. After all, the formula was based on quantitative data and mathematical techniques, and also offered very rich pickings, so what could possibly go wrong? With the pseudo-mathematics of the credit crunch, the Cartesian dream has turned into a nightmare.

The credit crunch is an indisputable example of mathematical science, animated by the Cartesian dream, getting it seriously wrong. This is a phenomenon for which the academic philosophy of science has provided absolutely no preparation. In the history of science, cases of scientists being wrong, usually in opposing an innovation, are explained in terms of their personal failings. The idea that The Scientific Method could betray us, is still, in practical terms, a heresy. Yet we must all become heretics if our reform of science and science advice is to succeed. We can accumulate some other cases of science, or science advice being significantly wrong, so as to strengthen our resolve. Most recently, we can think of 'the markets', 'cholesterol', 'germs', and 'clean, safe and cheap (nuclear power)', as examples of simplistic thinking that did not deliver the promised security or success. Each in its own way, they combined the Cartesian vision with a sort of American pragmatism: we don't have to solve problems, we just beat them to death.

In each of the cases I mentioned above, the scientists included people of the greatest talent, dedication and integrity. Those failures were much more interesting, and significant for policy, than those where scientists were simply incompetent or corrupt. Although the image of science has always been of a

form of knowing that was fundamentally opposed to fantasy, in these cases the element of fantasy has been strong and some times determining. I have a model to explain how fantasy can be destructive; and I believe that it can help us to understand how the Cartesian fantasy can harm. I call it the 'Eye-rack Syndrome', where 'Eye-rack' is the American pronunciation of Iraq. We saw it happening there in full view, and so we can list its phases. First comes fantasy; then denial and mendacity to maintain the fantasy in the face of contrary experience; then comes corruption as reality-testing is degraded and then lost; with that is incompetence, since quality-assurance becomes impossible; following on that is failure and perhaps disaster. In the case of Iraq, the fantasy was the American myth, that the world is divided into the good guys who love America and the bad guys who don't; and that peace and happiness are achieved by getting rid of the bad guys and replacing them with the good. The possibility that pulverising a civil society might produce chaos and resentment has no place in this comforting scheme. After all, it worked well enough with the Indians.

We are a long way from Iraq, morally and politically, when we consider the Cartesian dream applied to the quantification of the world. Yet, the pattern is there. It is easiest to see in the case of neoclassical economics, which reigned as the king of the social sciences until the crisis came, and the discipline that deified 'the markets' could neither predict nor explain the disaster. In the case of this science as others, 'corruption' need not, indeed usually does not, carry its usual vulgar meaning of taking cash for evil acts. Rather, it has its broader sense, referring to the degeneration of standards, to what has been called 'sloppy', 'shoddy' or 'sleazy' science, that is practiced or condoned by people who can (and sometimes do) see it for what it is. The classic image of science, implicitly assumed and purveyed by philosophers, publicists and teachers alike, accords no reality to this phenomenon. And so those specialists, and their audiences, are engaged in what is a group-level Denial, one which does not need any particular individual to be aware of its presence. And when an uncomfortable reality does obtrude, it can be safely relegated to the sphere of informal complaining, or gossip, in a sort of double-think operation among practitioners. The real historic significance of the present wave of criticisms of research science, including Ioannides and *The Economist*, is that they have forced these realities into the public sphere, attacking the Denial, and thereby threatening the previous official consensual fantasy of the perfection of Cartesian science.

With this model, taken from the politics of fantasy-led State violence, I hope to help to explain the interesting nonsense of the CO2 emissions from the Democratic Republic of Congo, and also that of the Gaussian probability distribution for the unmeasurable products of the 'quants'. My conclusion is that this sort of thing will continue to happen, regardless of all the well-intentioned movements for reform and rejuvenation of science, so long as science is in thrall to the Cartesian fantasies. For the Congo, as in all the other examples of pseudo-precision, the fantasy is the faith that for every imagined quantity there is a linear measure, which itself is unique and precise. That quantity may be one where only the traditional 'primary qualities' (length, time, etc.) are involved; then the reduction to a single linear *numeraire* may be plausible. But now we also quantify 'secondary qualities', including affect and emotional states,

producing numbers with great precision and no gauges of uncertainty. This fantasy is compounded by a special faith, that policy relevant science has the duty and also the possibility of delivering a unique, definitive, consensual answer to all questions. In this way the realm of values and complexity is tamed by objectivity and simplicity, just as Descartes saw in his vision so long ago.

Now we can see how the Eye-rack Syndrome can operate, not in a neat temporal sequence as in the original case, but rather in the policy context where there is a set of drivers and tendencies in a complex system of belief and practice. The dilemma is very real; policy demands simplicity in its imperatives, so that they can be defined and implemented. What would life be like for policymakers if they told the public that their crisp numbers used for defining implementation are actually produced from input data that is largely mush? The question could even be raised, why these policies rather than others? If the quantitative facts are really swamped by their uncertainties, perhaps the whole issue should be reconsidered. So it is much easier for a community of policymakers to adopt a scientific variant of Plato's Noble Lie, and to convince the public and oneself that these new guardians of welfare base their rule on simple, verifiable truth. For maintaining that myth, the supporting data must be cleansed of their impurities, and presented as definitive. And all this reasoning is conducted simultaneously at several levels, some semi-conscious, some informal and clandestine, and some systemic and implicit. It would involve a gross oversimplification, to accuse anyone of deliberate massaging of data. In a well run system, noone need know that they are corrupt. And the strategy may well be successful for an indefinite time, so long as the policies are not flagrantly at variance with the raw reality out there.

Some times, as in the case of the Gaussian copula, nemesis is not long in arriving. The Denial extended over the whole world where mathematical economics collides with reality. As to the corruption, the handful of technically competent critics identified how elementary principles of application of models were constantly violated in the practice of the quants. [Willmott, Tett] Gross incompetence was universal, and collapse came very quickly, as soon as 'the markets' began to behave badly. [Ravetz] In my other case, that of the pseudo-precise emissions data, the Eye-rack cycle is not so neat. The faith in numbers is so basic to our Cartesian world-view, and the misuse and abuse of quantitative information so pervasive, that the whole system of production and use of evidence for policy has evolved so as to tacitly accommodate such low-quality information. Noone in the Democratic Republic of Congo will make a crucial decision based on that pseudo-precise number. However, if such spurious information becomes decisive in determining broad policies, then it could happen that in the long run they will blatantly fail to relate to reality and so ultimately be rejected. Of course, that run might become very long, as we know from the cases I mentioned above, particularly 'cholesterol' and 'germs'.

Even in the world of basic research, it is becoming acknowledged that the corruption of quality is a serious social problem of contemporary scientific knowledge, not restricted to the occasional blatant fraud or widespread minor fiddling. Our colleagues Saltelli and Giampietro have demonstrated that the

problem is general, serious and dangerous to science. Although, as *Nature* reports, public trust in science, at least in the UK, is still quite high, that report shows an awareness that this situation is fragile. People would not suddenly abandon their faith in scientists, but a sequence of scandals could erode it seriously and irreversibly. I believe that the Editor of *Nature* would not consider it an outlandish suggestion, that science should now look to the Roman Catholic Church and be warned.

The problem of the provision of scientific information for policy thus relates back to the general problem of quality in science. If the situation is indeed serious, then what would be the elements of a solution? It helps to see it as a problem of corruption, although not at all a corruption of the conventional sort. Rather, as I have said, it relates to standards and practices, and can even affect scientific work that is intended to be responsible. There is just now a very important movement to restore research integrity, with much discussion of the various causes of the present situation. There are doubtless all sorts of pressures on researchers, owing to the sudden deceleration of the centuries-long exponential growth of scientific effort. Also, there are the pressures from the social institutions related to science, and (in England at least) a sudden awareness of large-scale corruptions and cover-ups of quite traditional sorts, in our most respected institutions.

For brevity in the context of this present discussion, I will offer a simple scheme for the causes of corruption in science, both basic and policy-relevant. I list three elements: metaphysics, morals and practice. Morals are the easiest point of entry. For some years now, the distinguished professor Sheila Jasanoff at Harvard, has been advocating a culture of 'humility' in science. This is much deeper than a simple change of attitude. It involves a change of expectations, which leads to metaphysics. By this I understand what sort of reality we believe ourselves to be living in, and hence what sort of knowledge we can have of it. This is where the Cartesian inheritance is crucial. For Descartes and Galileo as well, the world is mathematical. Only the measurable 'primary qualities' are real; the world of qualities exists purely in our perceptions. Each in his own way was committed to a programme of analysing and then controlling that qualitative world through a quantitative analysis. And part of that commitment was the belief that in that way, and only in that way, lies Truth about nature. Now, in what ways, and to what extent, that mathematical world must be enriched in order to be the basis of a healthy natural science, is a big question indeed, far beyond the scope of this lecture or this workshop. Contemporary physics has in various ways, opened the door to that quest. We can say that the concept 'systems' gets us out of the dogmatic reductionism of the past, without embroiling us in debates on enhanced realities of some sort or other.

Let me now pass to practice. For this we have a firm basis in the threefold foundations of quality assurance that Saltelli and Giampietro have proposed. As they say, evidence based policy has to be replaced by robust policy, where robustness is tested with respect to (1) feasibility (compatibility with processes outside human control); (2) viability (compatibility with processes under human control, in relation to both the economic and technical dimensions), and (3)

desirability domain (compatibility with a plurality of normative considerations relevant to a plurality of actors). We notice that all three criteria relate to context, either natural, socio-technical, or ethical, and so they are definitely post-normal. Complementary to these are the criteria relating to the internal sources of the evidence. For these we have NUSAP itself and the NUSAP tradition as developed by Jeroen van der Sluijs and colleagues at RIVM/NPL with the 'Guidance'. I can imagine a checklist to guide scrutiny of scientific projects in the policy relevant fields, whereby the Extended Peer Community can come into its own. For that, we can articulate a cycle of scientific inquiry, with phases along these lines: Issue, Policy, Project, Problem, Persons, Procedures, Product, Quality, Property, Publicity, Outcomes and Consequences. Of course, there will be connections and feedbacks among the phases, since we are dealing with a complex system. For all that highly structured intervention, I could recommend using some of the visual representations and aids developed by Robert Horn. And finally, if all this remedial activity helps to develop awareness of the need for a new arithmetical language to express uncertainty and quality (the rather neglected N of NUSAP) I will be very pleased to offer my ideas.

Our workshop will be challenging, and quite possibly it will be historic. My thanks again to the European Commission and to the Joint Research Centre for their support over the years and especially for this event.