



Speech to the Institute of Public Administration New Zealand (IPANZ) '*Communicating and using evidence in policy formation: the use and misuse of science*'

Te Papa Tongarewa, the Museum of New Zealand

21 February 2013

The relationship between evidence and policy is one of the most prominent topics in the conversations I have with my counterparts overseas. Indeed globally one sees the growing discussion of the importance of ensuring effective incorporation of objective evidence, science and data into the policy process. There are many reasons why.

Fundamentally it should be self-evident that effective decision-making requires good policy advice, and that in turn depends on the optimum use of objective evidence. But beyond that there is the increasing recognition that many global, regional and local risks have a major scientific and or technological component to either their origin or solution, and the problems that citizens want their politicians to solve are increasingly complex or, in that unfortunate jargon, 'wicked'. A characteristic of these problems is that the demanded solutions are urgent, the problems are complex with many unknowns and that there can often be pre-emptive and premature conflation of values with uncertain knowledge in the policy process. Many sociological, environmental and indeed economic challenges have these characteristics.

But we do not live in a technocratic society and the relationship between evidence and policy formation is neither linear nor unidirectional. It is rare indeed that any single piece of data is sufficient for a policy shift. Policy formation is a complex process, in which many factors other than objective evidence are properly brought to bear including the weighing up of many trade-offs. In such circumstances, it can be argued that the role of evidence in our kind of society is to 'nudge' this complex policy environment in certain directions.

It is broadly accepted that policy decisions are based on trade-offs across values domains such as fiscal priorities and affordability, public opinion, political ideology and electoral considerations. However, effective policy formation must also incorporate the use of objective evidence. Without such, the options and the implications of various policy initiatives cannot be measured against fact and when this happens judgment can only be on the basis of anecdote, dogma and belief.

Despite the best of intents, there is inevitably the potential for some cultural and attitudinal tension between those engaged in policy advice and those engaged in science (including social science and engineering). Because of these cultural divides, there are quite different understandings of policy development. Scientists tend to overestimate the utility of what they know and policy makers underestimate what they do not know. This can lead to a shallowness of policy argument, a promotion of political polemic and a narrowing of options. In some countries, think tanks, both public and private, are used to break down these barriers; but we have perhaps not the scale and certainly have not had the culture of such an

approach. More broadly, we are seeing many participatory democracies give more focus on finding ways to improve the use of objective evidence in the processes of policy formation, implementation, and evaluation.

But, because of the dominant role of public opinion, anecdote and political process in policy formation, there is a problem – misunderstood or misconstrued evidence can intentionally or unintentionally warp policy making. This has two implications: First, how society obtains and understands scientific and technical knowledge is critical to a well-performing participatory democracy, and second, in discussing the role of evidence in policy formation we also need to be aware of the consequences of poor scientific understanding and communication.

Too often a piece of science is misunderstood, misused, overstated or downplayed – sometimes something is presented as established science when it is not, other times it does not suit advocates to accept the prevailing consensus of science. A scientific position can be established even when all the details may never be resolved or there is still debate over some details. Classic examples include the Darwinian theory of evolution and understanding of the origins of the Christchurch earthquake.

Currently two matters are causing me particular concern.

The first has been the increasing trend for the complex nature of science to be misunderstood or even misused as can sometimes be seen when partisan politics engages with science. This can lead to the – sometimes rhetorically convenient – contention that you can find a scientist to support any given position. We have seen this in debates over climate, genetic modification, fracking and food safety. Scientific consensus is unlike social consensus – it is not a matter of the loudest voice or negotiated compromise. It is a process by which the expert community examines the currently available evidence and incorporates it into an understanding that integrates what we know and at the same time acknowledges what we do not know. The very nature of scientific observation means that results can be variable, and the process of scientific consensus addresses this problem. But the political and media processes can find it convenient to take another approach. But equally the scientific community can be at fault and I will return to this presently.

The second and related challenge is that of science being wrongly used as a proxy for a debate over values. This may occur consciously or unconsciously. It is obviously psychologically easier in some situations to say that the science is not settled, and then proceed to enter complex discussions based only on strong values components. Much of the climate change debate has used science as a proxy when the real and valid debate is over intergenerational equity. We have seen the same dimension in play in debates over new technologies such as genetic modification. And we have some scientists, who are human and have their passions, play into and exploit that debate.

Perhaps we should start by understanding the nature of science and its base for providing objective evidence. Science is an organised process for obtaining new knowledge – it is not just a collection of facts about how far it is to Jupiter or the number of cranial nerves in the dogfish. The scientific processes, for there are several, involve observation, generating hypotheses, gathering evidence to support or refute them, testing hypotheses by seeing whether they predict new facts, and modifying the hypotheses as needed to accommodate new evidence.

Critically, there must be a high level of integrity in analysing and interpreting data. This is protected to a considerable extent by two important foundations. These are replication and expert peer review by other qualified individuals so as to maintain quality and consistency. For these reasons, publication of a peer-reviewed paper in a reputable journal is considered a generally reliable – but not infallible – mark of scientific quality. So perhaps the first question for policy makers when interpreting a scientific claim should be – has it been peer reviewed?

Irrespective of this we must acknowledge that science is a human activity and scientists are subject to the same career and societal pressures as in any other profession. Just because a paper is published does not mean it is not flawed. Equally, scientists have their own personal values that may affect their interpretation of their work and the advice they give. They can become passionate and biased advocates rather than knowledge brokers and sometimes it is not clear what role they are playing.

Sometimes this can lead to tragic consequences – arguably the most extreme example was the Wakefield affair when a flawed and probably fraudulent study, published in a very reputable journal and subsequently widely publicised, cast doubt on the safety of the widely used MMR childhood vaccine. The resulting decrease in vaccination rates led to outbreaks of disease, causing deaths and permanent disability. While such examples are exceptional, they highlight the danger of reliance on single or extreme studies. In the case of the MMR vaccine, echoes still persist and are seized on by anti-vaccine advocates.

There are many technical issues involved in interpreting science.

There can simply be technical problems around study design or implementation that can cause problems that only expert commentary or investigation can detect. We saw an example in the announcement in 2011 of neutrinos that appeared to travel faster than light – apparently breaking a fundamental law of physics. The well-publicised paper caused a storm of criticism and commentary, and in trying to replicate their results the researchers found a loose wire in their apparatus that affected their timing measurements.

There can be variability in reported results and the conclusions reached about the same problem depending on many factors but most often the experimental design or simple natural variation and small sample size. There are many examples of where there is confusion – for example, does drug education in high schools prevent drug use or not? Exact replication can be difficult in policy-relevant social studies and indeed in biological and medical studies, and even assuming the same general approach was used, there can be an enormous variation in the claimed results because of differences in study populations and the details of the intervention.

Because of this variation, the danger of cherry picking is real where an advocate or politician will emphasize one particular study and its result because of a conscious or unconscious prior bias. But scientists can also be at fault; For many reasons too many scientists, aided and abetted by the communications departments of their institutions, are tempted to overstate the implications of their work, be they cures for cancer or pronouncements on the human condition.

These issues are arising more frequently because the nature of questions that twenty first century science now engages in are more complex and relate to social, environmental and human matters. This can lead to widespread confusion and when this happens, belief and dogma can become the basis for decision-making with the risks that then follow on.

Science has developed ways to address the many issues of variability which I will not go into tonight other than to mention the concepts of proper experimental design (so that the study has the statistical power to show what it is intended to show), *a priori* hypothesis setting (to avoid 'false discovery' due to chance) and meta-analysis (which is a way of combining evidence from different studies).

There is another fundamental problem that is too often misused or misunderstood. This is centred on the different meanings of association and causation. Such misinterpretation can occur in many social and medical issues. A current example is whether there is a causal relationship between adolescent television watching and adolescent behaviour. Some peoples have claimed that watching sexual content encourages early sexuality but equally those who have early sexuality may be more interested in such programmes. There can be many other sociological confounders in such studies.

As another example: the apparent incidence of child abuse is increasing. Is this real or are we just getting better at detecting and reporting child abuse? If it is real, then can we relate it to changing social circumstances simply because over the same timeframe family structures have changed? To be extreme and illustrate the point, it could equally be the result of greater use of cell phones that has occurred in the same time frame? The former may be more plausible, but the association does not prove causation. More complex analyses are needed. But in the end it can only be experiments or interventions that can prove associations are causations. In many cases, however, such studies are impractical or unrealistic. Here the choice to intervene must be based on only a plausible association because there is no other way to progress. In those cases, evaluation is critical – otherwise, investment may continue in a meaningless, even if apparently well justified, programme. However some countries, and the UK in particular, are taking the lead in thinking about randomised trials as part of social policy development. Clearly this is an approach that despite its political difficulties is in the long-term interests of any country.

But even if factor A does cause consequence B, there is a much bigger issue which is frequently forgotten and yet is key to policy formation. Scientific assessment of a result must consider not only if factor A *causes* consequence B but *how much of a change* it causes. So when looking at risk or benefit, we must think about the absolute rather than simply proportionate or relative effects. The media ignores this so often when it talks in percentage terms – a 100% increase in disease risk may still negligible depending on the base rate. This is an issue we often face in thinking about the effects of environmental chemicals, and we have recently seen in New Zealand an important example of how the minimal risk arising from one particular chemical residue in milk could have been better communicated.

The misuse of science in the public domain by politician or advocate can engender mistrust in the scientific enterprise. This is a real concern to me. The scientific community has to do much to improve its behaviour and understand the difference between brokerage and advocacy. But equally society will better served if science is not miscommunicated and not misused in advocacy or in policy formation. We live in a democracy and values will always and should be the final arbiter of decisions that are made. But values are formed and moulded by what we know and what we think. We can also bias what we learn from knowledge from our prior beliefs. But ultimately science is the only process we have to gather hopefully reliable information about our world, our society and our environment. It should therefore be seen as an essential input for all policy formation. But it can only do that if it is honestly represented and honestly used.

The challenges of the twenty first century are many. These include the obvious issues of climate change, sustainable economic growth, food, water and energy security as well as the more subtle issues arising from greater urbanisation, changed ways of communicating and the altered nature of our society. In the face of such challenges, science and technology will be essential to navigating a productive and safe path for the next generations. If science is not used and communicated in a way that is appropriate, we risk sailing into dangerous waters.

Ultimately, policy formation is about trade-offs in resource allocation. If dollars are to be spent on an intervention, the trade-off is that dollars are not available for another intervention. That is the basic reality of policy formation. Yet political advocacy played out in the media and elsewhere often obscures this reality.

We have seen this debate come to the fore in the discussions over class size versus other interventions, where some have argued to reduce class size for better outcomes and yet others have suggested that the effect of class size is trivial compared to other ways of enhancing pupil outcomes. This is a question clearly amenable to empirical research but other factors affect the decision-making process. The related issue here is to be clear about what outcomes of education are meaningful – is it student happiness, formal school performance or should we be looking at employment potential or progress through subsequent educational experience or should we even be looking long-term at variables related to integration into society (such as employment history, earnings or stable relationships)? The problem is that the last few measures may be what really matter, and that examination performance may be at best a surrogate measure of uncertain quality for predicting societal success. This example highlights the value of a more sophisticated and informed discussion that will improve outcomes for children and for society.

Opponents of the introduction of some new technologies often demand absolute proof of its safety. Some philosophers of science have argued that this is not formally possible. Similarly, proving that any technology has no adverse impacts can never be shown to be true, but can only be proven not to be true. Observations can show no adverse impacts, but this does not rule out that some other set of observations in the future may show some such effect. Instead the only rational approach must be one of risk and hazard assessment and adjudicating on whether the technology can be managed appropriately or not. Otherwise nothing new would ever be introduced.

Karl Popper famously gave the example of the statement ‘all swans are white’ to illustrate this point – to prove this statement would require laborious examination of the world’s entire population of swans, but to disprove the statement requires only discovery of a single black swan. We might not wish to forego the benefits of a new technology until, to continue the metaphor, all swans have been examined, but at the same time we should be alert to the appearance of a black swan.

Nicholas Taleb has extended the ‘black swan’ metaphor to consider the risks of very rare but high-impact events. Often we are more concerned about the possibility of a very rare event which has high impact – appropriately, meteors are in the news this week – when in reality more common events of greater familiarity have greater risk. Statistically, driving a car is much more dangerous than air travel, but people worry much more about air crashes than they do about their daily car journey.

In any intervention, there is also the potential for spill-over or ‘side’ effects which may be positive or negative – and the way these are assessed also requires careful approaches. If the study is not designed to look for these they may remain unknown. The issue of side

effects is critical to the policy maker – they need to consider what are the good and bad side effects of any intervention and avoid ‘unintended consequences’. Again there are ways to design interventional monitoring to look for these and assess how important they are.

So now let me return to the issue of the use of evidence and policy formation.

I have argued that the process of policy formation is improved if evidence is used in a generally value-free manner and only then should the various values-laden domains such as public opinion, fiscal priorities, diplomatic concerns and electoral considerations be overlaid upon knowledge. When the science itself is presented in a values-laden way it is compromised and loses any claim to a privileged place in the path to policy formation. Conversely, the failure to use evidence properly can lead to decision making which is less likely to produce effective and efficient outcomes.

But beyond the obvious issues of ethics, science is never absolutely value-free. The key values domains to consider are, first, expert judgements over the quality and sufficiency of data and, second, the limits of the data. There are nearly always inferential gaps between what is known and the decisions that are implied by the knowledge. The gaps and uncertainties must be acknowledged. If that is properly done then science advice can be delivered in an effectively value-free way. I have argued that the science advisor must generally act to present science in a manner that is not based on advocacy but is delivered by ‘honest brokerage’ to the policy maker. It is for the policy maker to overlay the other critical domains of policy formation.

From the policy perspective such an inferential gap is a challenge in several ways. Firstly, the lack of certainty can be used as an argument to avoid action and this can create extreme positions where the default becomes inaction. Secondly, because most policy decisions have to be made in the absence of full data and/or on the basis of uncertain evidence, the significance of the inferential gap between what is known and the policy actions that arise need to be clearly described and understood. This should inform evaluation processes both before and after policy decisions are made. From my perspective, advice needs to be proffered in a way that accepts both the presence of gaps in knowledge and the role of other domains in the decision process.

However, other forms of knowledge and apparent evidence exist and are in play in the policy process. I am not sure who first said it but ‘the plural of anecdote is not data’ and this highlights a key point. Anecdotes are often misleading, yet in the media and political process they are highly influential. Other informal ‘knowledge’ comes from spiritual, cultural and religious beliefs. Clearly such knowledge is not of the same nature or, in rational terms, have the same validity as scientific knowledge, but belief systems clearly have major impact through the values domains on the policy and political process. People’s values and beliefs of course matter hugely, but it is my strong view that such sources of ‘knowledge’ should not be equated to scientific knowledge. Beyond the objective knowledge base, it is for the policy and political process to weigh all other such inputs in the decisions that a democracy requires.

Many political decisions must be made in the absence of certainty of outcomes, no matter how intense the political conversation. The distinctive role of objective evidence is to help to reduce this uncertainty and therefore make the assessment of the options more informed. Broadly speaking there are few situations where science alone can provide complete answers to policy dilemmas; indeed in many cases it is science and technology that brings forward issues (sometimes unwelcome) that politicians must address.

And because ministers must often make policy choices in the absence of data or with very incomplete data, the role of objective evaluation *ab initio* of new policy initiatives becomes particularly important, albeit all too rarely employed. The value proposition of formal programme evaluation is often underestimated.

A worrying feature of the New Zealand science system is that, compared to other participatory democracies, there is a relative lack of process and investment surrounding the development of objective evidence to support policy formation. Our public science funding bodies have for at least the last 20 years not seen such policy-related research as a priority, and there is, in my view, insufficient connection between central agency needs and the scientific enterprise.

Similarly, the quality of policy programme assessment and evaluation is often not rigorous. Such scrutiny can be compromised or biased by agencies not wanting to embarrass the owners of a political decision. The evaluation process can be seen as unnecessary, especially where rhetoric has led to a strong political position. In general the understanding of the components of programme evaluation is weak across many agencies. The concept of controlled trials in public policy implementation is becoming well accepted in other jurisdictions but, perhaps given our short electoral cycles, this has not been generally accepted in New Zealand.

It can be argued that the issues I have raised are particularly acute in a small country such as New Zealand. Inevitably we have a less sophisticated system of connectivity between politicians, policy makers, lobbyists and the media. This, combined with the pressures created by a very short electoral cycle, results in greater potential for bias and exclusion of consideration of the evidence.

Some of the most intense debates in New Zealand, and indeed across western society, in coming years are going to be about the incorporation of new technologies (e.g. synthetic biology, neural implants, regenerative medicine). Similarly there will be issues around technologies to manage the balance between environmental protection and the need to enhance the economy. In such cases the science-policy nexus must assist the public and politicians to understand the risks and manage the opportunities and risks (a complex concept in itself) to create trade-offs.

Over the last year my Office has conducted a survey of attitudes to, and uses of objective data, across many of our ministries. We are still analysing the survey to prepare a report to the Prime Minister but the range of attitudes and practices is surprisingly disparate. There are examples of very good practice, but worryingly there are some attitudes at quite senior levels within our public service that suggest that there is a lack of understanding of the role of objective knowledge and expertise and of ensuring quality of inputs in developing policy advice.

Scientific input to policy can occur in a number of ways.

- Policy makers may think they have the competence to scan the literature and interpret the science. However, depending on the domain, there is grave risk that non-expert assessment of a complex literature can lead to cherry picking – that is, finding something to bolster a pre-formed bias.
- Policy makers may identify a knowledge need and go to a known expert or group of experts for advice. The quality of that interaction will depend on the nature of the question and the understanding of both the agency and the scientist of the science-

policy interface. Protocols have been developed elsewhere to clarify these expectations but are largely absent from our agencies.

- Policy makers and politicians may be lobbied by scientists either on issues related to science policy, or where scientists engage as advocates for particular causes. The role of academies such as the Royal Society of New Zealand in moderating such issues can be important.
- Policy makers may contract a piece of research either internally or externally. However, unless the contractors are skilled in such research procurement then the contracted agent, the research design and the quality of the resulting insights may all be at risk.
- Policy makers may invite experts onto advisory committees. This is to be applauded. However, if the required input really is that of scientific advice then only committees comprised of scientists can give such advice and this should come to the policy maker in an unfettered form. The alternative is a committee where other interests, such as those of end-users, will be conflated with scientific advice and objectivity is lost. To ensure independence and objectivity, scientific advisory committees require formal protocols of membership and operation. However, this does not mean that scientists do not have valuable contributions to make to other forms of advisory committees, but such committees alone cannot have the role of providing objective scientific advice.
- Departments (for example Conservation) may have internal scientific units that conduct science and research and these may be the primary source of broad scientific advice. The issue then becomes one of quality control.
- Departments may establish independent departmental science advisors with defined terms of reference. This system has been most well developed in the UK, where quality control of other inputs is an important component of their role. The Ministry of Primary Industries has recently established such a role here and I regard this as an important step forward.
- In some countries Parliament itself can foster the interaction. The Parliamentary Office for Science and Technology (POST) offers non-partisan advice on scientific matters to both UK Houses of Parliament. Parliamentary Select Committees (particularly where there are bicameral systems) focus on the quality and nature of scientific advice or conduct expert enquiries.

Let me suggest some possible ways to enhance the use of science and objective evidence in New Zealand's policy development framework

Both the UK Office of the Chief Scientist and the US Office of Science and Technology Policy have established protocols for obtaining independent scientific advice, either from individuals or from scientific advisory committees. Importantly, principles have been established with the goal of trying to ensure that such advice is free from bias or selective filtering. Given the variable state of the use of scientific advice and the limited and uneven appreciation of both its value and its limitations, such protocols need to be applied across the New Zealand policy framework. This is made more difficult because of the small corpus of expertise within New Zealand and thus the need for expert oversight becomes more critical.

My role was established in 2009, modifying and developing the model from that used in other jurisdictions. A key feature is its independence. This strengthens the neutrality of the role and gives focus to its broader functions of science communication, advice on evidence in policy formation, science diplomacy and specific scientific advice. These functions together, combined with a formal role in risk assessment and foresighting, are those that other jurisdictions increasingly see as critical to the role of a centrally positioned governmental science advisor.

I see considerable value in the creation of a community of science advisors each with specific terms of reference appropriate to their Ministry such as recently occurred in MPI or group of Ministries. This would, in my view, add great value to the quality of policy formation in New Zealand. Each appointment would call for a wise and broadly experienced scientist in the appropriate discipline. It is not appropriate that such roles be filled by people without a track record of achievement in scientific research. But science advice is not a matter of being expert in every technical field in which advice is sought – rather it is having the knowledge and insight to know where to go for information and then interpreting it appropriately (including assessing limitations and possible biases) for use by the policy community.

A particular set of issues relates to consideration of the scientific approach to risk assessment and risk management. The recent Global Risk Report by the World Economic Forum indicated that most major risks have strong scientific or technological component. Many of the challenges we face going ahead are balancing the use of current or new technologies against the risks of such technologies. This is an area where the formal analysis of risk and its implications for risk management can often diverge significantly from broader perceptions of risk. Political, scientific and public understandings of risk differ widely. It is important that decisions are made with an objective level of appreciation of the actual risk as well as the public's understanding. In some areas such as earthquake risk management (e.g. building codes), managing volcanic ash clouds, and other natural disasters, there is a primary role for the formal engagement of the scientific risk community. The science advisor has a core role in such processes.

Taxpayer funds are used in three ways to support research for policy formation: firstly, agencies undertake research intramurally, secondly, they contract research extramurally, and thirdly some of the contestable research undertaken by the Crown Research Institutes and universities as a matter of course relates directly to the policy agenda. The first two of these activities do not necessarily meet the standards that research funded through contestable processes is required to meet. Internationally it is accepted that governments want academia to engage in research that has policy implications. However, the manner in which the priorities for such research might be provided to the funding agencies is capricious and uneven. Again the departmental science advisor could have a core role in addressing these issues – indeed that would be central to her brief.

Part of improving the use of government funds is also to improve the focus and commitment to programme evaluation. Ministers should expect and demand that more programmes are subject to efficacy evaluation, that funds are allocated for that purpose, and that reviews consider not only new programmes as they are rolled out, but where possible current programmes. There should be no political embarrassment in acknowledging that the impact of a new programme is not known and must be evaluated. A good example of this approach being taken is in the youth mental health programmes announced last year. There should also be a greater willingness to consider pilot programmes and undertake careful evaluation of success factors in the case it is decided to go to scale.

The Government is in the process of establishing a Social Policy Research and Evaluation Unit within the Families Commission, with a scientific advisory board. This will provide an autonomous unit with expertise in social science research and in ensuing evaluation – it is hope that it could support activities across multiple ministries. While this unit is in its early days, it will be critical that the Commissioners ensure that its standards and mode of operation are of the highest quality. It must develop specific skills in programme evaluation such that Ministers will wish to encourage agencies to take advantage of its expertise.

The other arm of policy-making and evaluation within a democracy is Parliament itself. Parliament may wish to consider whether some of its activities could be better informed and whether, through the Select Committee process, greater focus could be given to the quality of evidence presented. This would require consideration of the same issues that have been discussed with respect to policy advice. This is more related to their role in auditing departmental performance rather than their management of submissions in relation to legislation. I am not arguing that evidence be seen as a political tool, rather it would be parliament taking on the role of ensuring quality of evidence coming forward. Departmental science advisors might have a role in assisting. The alternative would be to develop a role akin to the UK Parliamentary Office of Science and Technology.

In making these comments it is important to note that strengthening and assuring the use of objective evidence in policy advice in no way weakens the political process. On the contrary, it tends to strengthen it. For in the end the political process is about making hard choices based on a range of complex options where there are inevitably trade-offs, spill-over benefits and costs. All of this occurs within a complex and unstable environment where human responses and decision-making are influenced by many factors other than objective knowledge.

Humans are characterised by the ability to create knowledge and use it. Technology helps us and challenges us. As we live in a more complex world, more and more complex decisions are required of our policy makers. Good use of objective evidence will help them meet that challenge better.

Thank you.

ENDS