THE CHARACTERISTICS OF RISK ASSESSMENT RESEARCH

RON JOHNSTON

Policy Division, Department of Science, Australia

Introduction

Risk has become big business. The civil nuclear power industry from the beginning was aware that it had to develop special methods for limiting and controlling the hazards produced by nuclear fission. These hazards have turned out to be far greater and more various than first conceived and as a result this industry is having to devote most of its efforts to issues of risk. Concern over the potential risks of pharmaceutical products, raised to public attention by the Thalidomide disaster, have led to companies in this field having to spend more than half their research budget on testing to meet safety standards and the research costs for a single successful drug are now estimated to be from U.S. $1–5 million. In the chemical industry risk investigation now accounts for 25–30% of all R & D costs involved in product development.

With this growth in concern over the expenditure on risk, there has developed a wide variety of types of risk research and study.[1] Risk assessment is the general rubric which has been applied to all these studies, which include the determination or 'measurement' of the objective risk of a particular hazard, evaluation of groups at risk and likely damage to them, management of risk through design, regulation, monitoring and training, and such specific techniques as risk impact analysis which attempts to provide a quantitative basis for comparison of risks, event and fault tree analysis for determining the topology of accident sequences and safety engineering.

Can any generalisations be made at this early stage about the state of risk assessment research? Perhaps not surprisingly, the field might best be characterised as fragmented, with very little contact of people or flow of ideas between the various sub-fields. Indeed the most obvious characteristic is one of conflict and lack of agreement between the protagonists of the various sub-fields, and even between researchers within a field.

There is not even a consensus over whether the central concept of risk can be operationalised and objectively measured. Scientists within the nuclear industry have to a large extent been responsible for the development of the technique of probabilistic analysis, by which, it is argued, the level of risk (expressed usually as the 'fatal accident frequency rate') acceptable to society can be quantified.[2] Thus experts can determine safety levels. By contrast, Lowrance emphasises that determining safety involves two different kinds of activities:

> *Measuring risk* – measuring the probability and severity of harm – is an empirical scientific activity;
> *Judging* safety – judging the acceptability of risks – is a normative political activity.[3]

Failure to appreciate this

> gives rise to the false expectation that scientists can *measure* whether something is safe. They cannot, of course because the methods of the physical and biological sciences can assess only the probabilities and consequences of events, not their value to people. Scientists are prepared principally to measure risk.[4]

Other commentators[5] are dubious whether even the measurement of risk can be left to experts:

> the numerical severity of risk is only one dimension of it, and is an insufficient parameter on which to base a comparison. Just because the risk of working in a particular industry is no greater than, say, that of being struck by lightning, is no reason for claiming that both are acceptable. As Jeremy Bentham pointed out, an evil is not justified by the mere existence of greater evils, it may be that the risk of being killed by lightning is not acepted, but tolerated only because of the expense of any remedial strategy. The problem of one-dimensionality is not solved merely by introducing a distinction between voluntary and involuntary risks, because there is no such thing as an involuntary risk: all risks can be ameliorated, although the cost of doing so may be prohibitive. If we can *choose* not to ameliorate a risk, it is voluntary.[6]

The organisers of this Workshop have directed our attention to 'risk assessment as an exemplary case of the interaction of science and politics in its historical, scientific and socio-political context'. There are, I think, good reasons for this. As sociologists, we are in a very advantageous position to observe the emergence of a new field, as evidenced by the trappings of institutionalisation such as conferences, publication of a growing number of articles and books, regular appearance in the public press and no doubt soon, a journal. A useful contribution can be made then to the sociological literature on specialty formation.[7]

Secondly, there can be little argument over the extent to which this research is constrained and directed by the political issues associated with risk determination and control. It is then, *par excellence*, a field in which organisational pressures and political objectives shape the selection, production and evaluation of scientific knowledge and its study can contribute to the development of more comprehensive theories of the interaction between political and intellectual aspects of the social reality which constitute science.[8]

Finally and most importantly, it is evident that risk assessment research has potentially much to contribute to the resolution of critical issues of our time. The civil nuclear power programme, the development and use of pharmaceuticals, pesticides, food additives, construction materials, all will be determined to some extent by the theories, concepts and data provided by the newly emerging risk assessment community. To be even more dramatic, risk assessment will influence the death, the life span and the quality of life of many people. There is almost none of the distance from public issues which characterises most scientific research, and therefore the responsibility is evident and immediate. Those who can offer an insight into the processes whereby knowledge contributes to the political perception and determination of issues, have a special responsibility to translate this essentially academic knowledge into a form in which it can be used to interpret and explain the forces at work in the development of risk assessment, and in particular to demonstrate the political assumption and limitations of such research.

In response to all these challenges, I wish to argue that the form of risk assessment research, its progress or lack thereof, its fragmentation and its political contextuality can be explained by three major characteristics of the field:

1) its goal-direction
2) its lack of maturity
3) its reliance on inappropriate models.

I have argued elsewhere that one of the distinctive features of contemporary science is that it is goal-directed, and only an analysis which incorporates this can hope to comprehend its form and content. Goals which are an

> over-arching objective or end which serves in part to direct one or more fields of research and of which individual, or groups of, scientists need not be explicitly aware [9]

explain the interaction between scientific, political, economic and social interests in the shaping of scientific knowledge. Immaturity is a state in which the value of the evidence, information and data being collected, and the conclusions of arguments based on them, are not widely agreed upon by members of the research community. The appropriateness of models in a particular field can only be assessed in terms of their fruitfulness in developing stable, consensual knowledge relevant to the subject matter of that field.

In order to explore each of these three propositions it will be necessary to examine risk assessment research in some detail. Given the variety and disparate character of the research conducted in this general area, as noted earlier, it will be necessary to focus on a selected sub-set of the total. Two fields which have been extensively developed, have attracted a significant number of adherents, and have made a distinct public impact are 'probabilistic analysis' and toxicology. We will examine each of these as test cases for the three theses outlined above. But first, a brief description of the two research specialties.

In probabilistic analysis, the method used to determine the probability of an accident begins with the construction of a logic tree leading from initiating events to final development. The logic structure consists entirely of 'and' and 'or' gates, causally connected. For example, if two open valves were in parallel, it would require a failure to both to close a fluid line, and the failures would be fed into an 'and' gate before the consequence (closure of the line) could be assumed to occur. Clearly, there is an enormous complex of such structures, (but apart from questions of continuous variables, which will be discussed later) the method in principle is able to cope with all possible sequences of events. Unfortunately, the number of interconnects soon becomes impossibly large in a real system leading to perhaps billions or more of possible logic tracks through the system from beginning to end. Because this is so, all the so-called different methods of analysis, or different methodologies, are simply alternate means of extracting from this intricate topology of sequences small and meaningful subsections which are tractable, and the application of differing levels of quantitative analysis to those subsets excised. For example, event-tree analysis consists in limiting the universe of discourse by beginning at one particular event, and tracing forward in causal sequence through those paths which derive from it, so that one considers all possible consequences (within reason) of a particular initiating event. Fault-tree analysis is exactly the same, except that the event provides a starting point from which one traces connections backward in time, thereby discovering all those earlier circumstances which can lead to the event under consideration.

The successful implementation of such a methodology rests heavily upon the availability of an adequate data base, effective computing tools, a sufficiently deep understanding of the detailed engineering of the system to permit construction of the logic trees and some logically sound procedure for limiting the universe of discourse.

Toxicology is the major science of toxicity. The main aims of toxicology have been described as:

i) The quantitative analysis of the effects of chemicals in biological tissue, using tools derived from statistics, analytical chemistry and biology.
ii) Identification of the biological mechanisms which produce the harmful effect when a chemical is introduced at a given site. This concern is allied to that of pharmacology and biochemistry.[10]

The concerns of practitioners of toxicology were, until the 1940s, predominantly with the quantification of toxic effects and the mechanisms of acute i.e. short term, poisoning, and in this they achieved a reasonable degree of success.

However in recent years, there has been growing controversy over the selection and interpretation of toxicological evidence and adequacy of toxicological theory. This has been fostered by the vastly increased use and dispersal of chemicals in the human environment, as well as other, more technical factors such as the increasing sensitivity of analytical techniques and the demonstration that chemicals could induce chromosome changes and thereby cancer.

Disagreements over basic toxicological data usually focus on a number of issues related to the design and interpretation of experiments. The time and cost of carrying out detailed toxicological investigations are considerable[11] and this means that in tests for toxicity and especially chronic toxicity a minimum number of animals tend to be used often at high dose levels in order to reduce the time and cost of testing. For similar reasons, chemicals are tested on a minimum number of different species. Consequently, there are frequent debates over the animal species, dose levels, animal numbers and data presentation techniques which should be used to supply a basis for risk assessment. Attempts to formulate 'acceptable' procedures have met with little success, as 'standard' tests in both short and long-term toxicity have been shown to produce inconclusive results..

Risk assessment as goal-directed research

Goals form an important, and thus far largely unconsidered, element of the cognitive and hence social structure which provide the milieu within which the scientist works. Goals may vary greatly in type, from highly mediated expressions of socio-economic and cultural context such as operate in physics and biology, to more direct expressions of non-scientific interest groups such as sectors of government, industry, or the public, as seen in the development and direction of fields such as computer science, geology and environmental science.

Particular goals may also differ in their 'level' of effect. Thus,

> the development of Kuhn's models suggest that cognitive structures operate at different levels and thus it appears reasonable to infer that goals, which to a large extent will express themselves through the cognitive structures, may also operate at different levels.[12]

High level goals form part of a tacit background knowledge internalised through socialisation which may not be consciously associated with many phases of scientific work but which nevertheless provide a powerful directing influence on research. They may also be formulated at the level of the specialty where they will take a more concete and explicit form, as, in the phage group example I have previously used, the determination of 'the mechanisms by which genetic information is transferred'.[13]

Three characteristic features or phases of the goal direction of scientific research have been identified:[14] goal establishment, goal mediation and goal evolution. Scientists, and those concerned with directing and influencing research operate according to goals which are:

1) Established as the result of social and political processes which involve dynamic interaction between interest groups both involving and excluding direct scientific interests and which may be directly or indirectly perceived by scientists.
2) Mediated by scientific, socio-economic and political considerations and expressed at varying levels of generality; these mediated versions may be expressed within 'official' statements of research programmes or they may be deeply embedded in the cognitive structure of the relevant sciences.
3) Dynamically linked to an evolving body or bodies of scientific knowledge in such a way that goals and science are only analytically separable; both cognitive and social aspects of research are directed and constrained by orientation to goals which are posited and potentially continually redefinable in terms of changing theory, techniques and socio-economic conditions.[15]

We will now examine how each of these features fit the two selected cases of risk assessment research – probabilistic analysis and toxicology.

Probabilistic analysis, or 'hazard tree' analysis has its origins in systems analysis. The metaphors of the electrical circuit with its possibilities of logical path analysis have been extended widely to manufacturing processes, architectural design, organisation management and even to social planning. The essence of the systems approach is that any complex event or process can be reduced to its components and the logical relations between those components clearly established. As a method of conceptualising and comprehending interconnected systems it has been highly effective. The added dimension in hazard tree analysis is that the probability of any particular system failure occurring i.e. each node point, must be either calculated on the basis of previous experience, or estimated.

Probabilistic analysis has been developed almost exclusively in relation to the civil nuclear industry, and largely by the government promotion agencies.[16] Thus, in the U.S. it was developed within the Atomic Energy Commission until succeeded, as far as regulation was concerned, by the Nuclear Regulatory Commission. In the U.K., it was the Safety and Reliability Directorate under Farmer at the Risley office of the U.K. Atomic Energy Authority which developed this approach to a high degree of sophistication.

It is evident that the development of this field of analysis, quite regardless of its validity, was not a consequence of pure scientists following their own personal curiosity. The majority of the scientists worked in government institutions,where the objective was one of the development and promotion of civil nuclear power and as a result their research would, at the very least, be constrained to fit within the broad mission of the agency. Secondly, it was not an intellectual, but rather a practical and political problem which generated the interest in probabilistic analysis. The economic development of civil nuclear power has from its earliest days rested critically on being able to demonstrate that under all reasonable circumstances the extraordinarily dangerous fissile material could be safely contained. More recently, as public scepticism and criticism has grown, the government regulation has become much tighter the need to be able to calculate and demonstrate that all conceivable hazards have been allowed for and cannot occur has added an even greater incentive, or pressure to develop techniques to assess the risks of nuclear reactors and nuclear fuel. As there is no human experience of many of the possible failure modes, this has led to the development of these new techniques for assessing the probability of 'hypothetical' events.

Thus the development of probabilistic analysis, and the form it has taken can be explained only if we take into account that the goals of this research were 'established as the result of social and political processes'. Let me hasten to emphasise, as this has frequently been the source of much confusion, that this does not directly imply anything about the quality or validity of research carried out under such goals. There is no *a priori* reason why goal-directed research should not have the validity of any other area of science; that issue must be resolved separately and for each particular case. But it does show how politically-based goals become incorporated into, even constitute, scientific fields.

Transformation of the high level goal of 'developing methods to calculate the safety of nuclear reactors' was necessary before it could become a practically useful guide to scientific work. The resultant concrete goal of developing hazard tree analysis emerged as a consequence of mediation by the existing state of knowledge of path analysis, of concepts such as the 'fatal accident frequency rate' (FAFR) and the data which had been collected to measure this for a variety of activities, and of techniques for estimation of hypothetical risks. Socio-economic forces such as the nature of the nuclear industry, the economics of generation of electricity by nuclear power and public concern had an equal effect, as we have seen, on producing the specific goal in its particular form.

Socio-economic mediation has played a particularly influential role in goal evolution, as public pressure and government directives have shifted the focus from one area of the field to another. Thus at one time the major developments were in the improvement of the statistical base including the validity of the 'square-root bounding model' used to estimate common cause failure probabilities, extending the data base and examining the validity of various kinds of distribution for fitting data, the limits of error associated with subjective probability and probability smoothing and the problem of propagation and multipication of error through a long series of calculations. At another time the major focus has been on the adequacy of the procedure for aggregating sequences or dealing with the existence of continuous variables which do not fit the dichotomous logic structures of hazard tree analysis.[17]

Turning now to toxicology, from the definition provided above it is apparent that it is a goal-directed research field; its goals are specifically stated as the identification and measurement of the effects of chemicals on biological tissues. Study of toxic materials has a long history but it is only in the twentieth century that it has become at all organised. To argue that toxicology is goal-directed may appear rather obvious though it should be remembered that it required considerable political pressure in the nineteenth century for governments to intervene in the market and ban or control the use of certain food additives.

However goal mediation and evolution have been quite pronounced in the field of toxicology. A quite different range of problems and pressures were presented to toxicology with the dramatic growth of the petrochemical industry and advances in organic synthesis. Enormous numbers of new chemical products started to enter the market; to give some indication there are now approximately 70,000 chemicals in routine use. Moreover as well as the clinical effects of massive poisoning, evidence began to appear of sub-clinical chronic effects of low level concentrations of materials which had been previously considered safe (or more precisely, not considered unsafe). These changes in the subject matter, together with advances in analytical techniques, improved understanding of human metabolism, an increasingly regulatory stance of government and public pressure have quite transformed the working goals of toxicology.

The growth of toxicology in response to these demands has occurred predominantly in the settings of industry and government. Most toxicologists are employed in large chemical firms, small specialist toxicological contract companies, or government laboratories concerned with the development of standard tests and provision of appropriate data. There are a relatively small number of toxiologists in universities and even their independence has, in certain cases, been questioned.[18] As for probabilistic analysis, the general objectives and viewpoint of the organisation for which the scientists work has shaped their research, influencing not just the evaluation of data, but also how, when and *if* it will be produced at all.[19]

As public, and hence governmental, concern has grown over safety issues toxicologists have found themselves pressed into government service as risk assessors, at the same time as debates over these issues have become increasingly public. Thus they have found themselves at the centre of a series of polarised conflicts. The conflicting professional and political roles toxicologists play have engendered grave doubts about the adequacy of their expertise. For example,

> the revelation that the regulation of certain drugs and pesticides was based on fraudulent data has confirmed the suspicion with which

> critics have regarded the quality of commercial and often secret data. Episodes such as this have created pressures on government to intervene in what was traditionally the preserve of the expert. Thus governments are prescribing what constitutes 'good laboratory practice' and endorsing particular experimental designs and interpretive principles to be used in risk assessment.[20]

One could hardly ask for a clearer demonstration of external socio-economic direction, not only of goals but also of practice and standards.

I would suggest the evidence clearly supports the argument that both probabilistic analysis and toxicology are strongly goal-directed fields and by extrapolation it may be reasonable to consider the whole field of risk assessment as goal directed. As I have already noted this does not imply directly any criticism of the quality of the knowledge produced within the field. Rather it shows how research in a particular field can incorporate economic and political values, which can itself be a target of criticism for quite different reasons.

The maturity of risk assessment research

As was already noted an immature field is one in which there is little consensus over fruitful lines of inquiry or even competent research. Ravetz[21] has produced the most penetrating characterisation and analysis of immaturity and in this section I will draw extensively upon it. In describing the state of maturity, he argues that the failure of most conclusions of arguments, and the evidence and data on which they are based, to survive

> even in the short run, is an indication that most of the work of investigating problems is vitiated by pitfalls, encountered sooner or later in the work. The results of research are generally weak, or even vacuous. This condition prevails even in fields where the leaders and their associates spare nothing in their endeavours; but the absence of a body of appropriate methods of inquiry nullifies their efforts. For it is through such methods, ranging from the techniques of production of data, to the judgements of adequacy on an argument, that pitfalls are identified, and ways around them are charted. Because of the subtlety and sophistication of scientific inquiry, these methods are a craft knowledge, built up by successful experience. But an ineffective or immature field has no such experience; and so the improvement of its methods is not a straightforward operation.
>
> The weaknesses in the social aspects of inquiry also contribute to the self-perpetuating condition of ineffectiveness. The mechanisms for the processing of results, and for the exercise of quality control, cannot

be stronger than the materials on which they operate. For social reasons it is necessary to give the formal authenticity of publication to masses of results which are very weak; and so the effective standards of quality cannot meet those of a matured field. Because of the rapid succession of separate schools, each with its own objects of inquiry and principles of method, there is little opportunity for results of potentially high quality to survive and become established as facts. And in this unstable and frequently false social situation, the mechanisms for the control of quality and the maintenance of scientific integrity at the highest levels do not exist.[22]

This description would appear to be appropriate for risk assessment, and probabilistic analysis and toxicology in particular, characterised as they are by fragmentation, lack of coherence of knowledge and consensus of researchers, and little evidence of the establishment of a stable and qualified community.

Naturally the leaders of these fields either deny their field's immaturity or claim that it can be easily rectified. In response,

Heroic attempts have been made to amass empirical data; to apply mathematical and computational tools for the production of information; to construct elaborate systems representing objects by symbols and manipulating them in a formal argument; and finally to develop methods and a methodology appropriate for the discipline. In each case, the attempt is to reproduce what is believed to be the crucial feature of an established science, where this is learned more from philosophers of science than from the successful practitioners themselves. But almost all these one-sided efforts fail utterly.[23]

Ravetz's argument is almost a caricature of the efforts of the hazard tree analysts. The issue of focusing on a presumed crucial feature of science will be examined in the next section.

How can a field gain maturity? There is no guarantee that this can be achieved. Only if the objects of inquiry can be conceptualised in a fruitful and stable form that allows manipulable elements which correspond to some feature of reality to be extracted can a start towards maturity be made.[24] There is little indication that this has been achieved yet in either probabilistic analysis or toxicology.

The pretence of maturity itself creates a range of problems which further inhibit the advance of the field:

Recruits are generally given no warning that their research work is likely to be very hazardous, and after some years of producing results which inexplicably fail to consolidate into facts, they can become demoralized; the real state of their field becomes a shameful secret. Nor do they have the security of knowing that their years of specialized training in the subject gives them a monopoly of practice in it, as in a learned

> profession or a matured science. Moreover, since the social mechanisms for quality control and direction in the field cannot function properly, the safeguards against the abuses of prestige are weakened, and the assessments of a lay audience, based on popularizations, can be of more practical importance in the politics of the field, than those of the community of experts. Under such conditions, the pretence of maintaining the social mechanisms appropriate to a matured discipline, and even more the task of improving their real condition, are rendered yet more difficult.[25]

The pattern of publication in toxicology reveals that most researchers are primarily attached to another discipline and contribute only for a short period of their professional careers to this 'marginal' field.[26] The extent of lay assessment of both fields, and risk assessment in general has already been indicated in the previous section.

The most serious problems arise when immature sciences are enlisted for the solution of practical problems. This inevitably leads to deception of self and others; for

> A discipline which is unable to establish facts even within its closed world of controlled experience is much less capable of genuinely drawing conclusions about the problems of a raw and unstable reality.[27]

How serious then must be these shortcomings for the two fields we have examined, born as they were directly to deal with practical problems? No opportunity existed to develop a consensus prior to the posing of such problems.

The consequences of this lack of independence are such that

> When an immature field takes on the task of expanding its research effort for the solution of some urgent practical problem, there will be a tedency for the outcome of its labours to be a weighty argument establishing the conclusions that its sponsors and its public wanted all along.[28]

Here the arguments of goal direction and immaturity converge to suggest that the formal content of risk assessment research may reflect the interests of its sponsors and sympathetic public and that such fields have no claim to the authority, or the ideology of neutrality, of science.

The models used in risk assessment research

There are features inherent in risks which make it an extremely difficult subject for development of sound knowledge. Ravetz has suggested that there are three basic contradictions of risks which render them

> unique among problems of an ostensibly experimental or scientific character . . . There are important senses in which risks are incalculable, unimaginable, and uncontrollable.[29]

For the purposes of this paper we are interested in the first of these. Setting aside those hazards of which we have a great deal of experience, such as motor vehicle or aeroplane accidents, risks are below the limit at which consistent calculations can be made with reliable quantitative estimates of probability. Whether we take the major hazards which are the concern of probabilistic analysis, where the probability of occurrence is extremely small, but where the consequences of that occurrence are catastrophic, or the quite different problem of chronic hazards with which toxicology deals, where the effect is so pervasive and hence difficult to isolate from all other factors at work, the problems are essentially 'trans-scientific', incapable of firm resolution by scientific methods.

There are no simple chains of inference, whether inductive or deductive, nor is there any obvious basis for either conformation or rejection. Hence the line of causal reasoning, which is the basis of theory construction and evaluation in so much of science, is simply not appropriate. Proponents and some philosophers have argued that probabilistic relationships are only a subset of causal ones. However I would suggest the degree of confusion and controversy between philosophers, mathematicians and physicists over the meaning and measurement of probability makes it insufficiently robust at the present time to carry the load of providing precise measures of risk within quantifiable limits of error.

On this basis, it might well be reasonable to argue that the model or paradigm science adopted by risk assessment research, namely physics, is quite inappropriate for its subject matter. That probabilistic analysis is modelled on physics should be self-evident. That toxicology is following the same blueprint may require a little more support. The argument is too lengthy to produce in detail here but the reliance on exposure-effect correlations, and such crude measures of toxicity as the median lethal dose (LD 50) reflect an attempt to establish universal causal relationships between material and effect and a model of combining precise experimentation with mathematical theory. For both fields, the object is one of providing precise quantification of risk.

It may be that a much more appropriate model for all of risk assessment research would be that of traditional engineering rather than physics. In this approach, while mathematics is used to provide