Pergamon

Marine Pollution Bulletin, Vol. 36, No. 12, pp. 939–950, 1998

Printed in Great Britain

PII: S0025-326X(98)00113-1

0025-326X/98/$19.00+0.00

VIEWPOINT

The Precautionary Principle: Protecting Against Failures of Scientific Method and Risk Assessment

D. SANTILLO*‡, R. L. STRINGER*, P. A. JOHNSTON* and J. TICKNER†

*Greenpeace Research Laboratories, Department of Biological Sciences, University of Exeter, Prince of Wales Road, Exeter EX4 4PS, UK

†Work Environment Program, University of Massachusetts Lowell, One University Avenue, Lowell, MA 01854, USA

The precautionary principle has been increasingly incorporated into national and international legislation to protect the environment from anthropogenic impacts. Suggestions for redefinition of the principle have been made (Gray and Bewers, 1996). Adopting these suggestions would entail using a risk-based approach to environmental protection rather than a truly precautionary approach. There remain enormous gaps in our understanding of the identity and of the individual or combined toxicity of chemicals released to the marine environment. This is also true of the actual doses involved in environmental exposures, whereas the most appropriate ways of determining ecosystem impacts have not yet been identified. Classical toxicological assumptions applied to dose–response relationships can no longer be regarded as universally applicable. Ecotoxicology, consequently, contains irresolvable indeterminacies. It is clear, therefore, that the continuing application of the precautionary principle as a paradigm for regulatory action, as a means of recognizing and accounting for limitations to scientific evidence, is an approach that is both scientifically defensible and capable of ensuring a high level of environmental protection. © **1998 Elsevier Science Ltd. All rights reserved**

Keywords: precautionary principle; environmental protection; legislation; policies; risks.

The Viewpoint article by Gray and Bewers (1996) represents an interesting, but misguided, contribution to the long-running debate surrounding the precautionary principle. The authors suggest that the application of the concept should be extended to address physical impacts rather than only those that are related to chemical inputs. They argue, also, that the definition of the precautionary principle should explicitly endorse a role for science. To be clear at the outset, application of the precautionary principle to non-chemical impacts does not require a redefinition of the principle itself. Further, there is obviously a role for scientific activity in marine environmental regulation, management and protection. Indeed, we would wholeheartedly concur with Gray and Bewers observation and the enthusiastic subsequent endorsement of this by Chapman (1997) that marine environmental protection activities require more science rather than less. There is a pressing need to increase our understanding of marine systems in order to formulate effective regulation, to implement it and subsequently to assess its effectiveness. This implies increased research effort and suitable channeling of adequate resources. It must be acknowledged, however, that in many cases, the existing scientific data base cannot be used to justify a particular course of action, or indeed to justify taking no action at all. In this context, a decision not to act in the face of uncertainty is as much a policy decision as taking precautionary action. In these situations, precaution is needed. This keystone principle has gained widespread acceptance internationally since the late 1980s.

Regulation and Definitions

It was recognition of the limitations of scientific knowledge that led originally to the formulation of an approach to environmental protection which was fundamentally precautionary in nature. Broadly speaking, such an approach recognizes limitations to scientific and technical information and promotes

‡Corresponding author.

regulatory action to prevent, or avoid the threat of, environmental harm before it has occurred, even in the absence of full evidence of a cause–effect relationship. In essence, the intention is to allow for incomplete data, uncertainty and indeterminacy to be taken into account in a meaningful way in order to allow decisions to be made where necessary. This intent is quite clear from *inter alia* the 1987 Ministerial Declaration on the North Sea. This declaration was almost exclusively directed at regulation of polluting substances from various sources and in it Ministers agreed:

To accept the principle of safeguarding the marine ecosystem of the North Sea by reducing polluting emissions of substances that are persistent, toxic and liable to bioaccumulate at source by the use of best available technology and other appropriate measures. This applies especially when there is reason to assume that certain damage or harmful effects on the living resources of the sea are likely to be caused by such substances, even where there is no scientific evidence to prove a causal link between emissions and effects (the 'principle of precautionary action') (MINDEC, 1987).

Gray and Bewers (1996) have taken this as a basic definition of the precautionary principle and suggest that the text be amended to take into account impacts resulting from other human activities. Their suggested definition was as follows:

To accept the principle of safeguarding the marine ecosystem of the North Sea by reducing emissions of hazardous substances at source and minimizing physical disturbance of marine habitats caused by human activities using appropriate technologies and measures. This applies to all human activities for which there exists a scientific basis for believing that damage to habitats or harmful effects on marine living resources are likely to result. Measures adopted should be based on pessimistic assumptions regarding uncertainties in the measurement and prediction of effects on the environment of the North Sea.

However, the Ministerial Declaration (MINDEC, 1987) did not attempt to define the precautionary principle, merely to implement it with respect to regulation of marine pollution in one geographical region. Gray and Bewers' proposals (Gray and Bewers, 1996) for alternative wording of the 1987 Ministerial Declaration reflect a number of areas in which precaution is necessary, but represent only a subset of its potential breadth of application. To accept their suggested wording as a full definition of the Precautionary Principle would be to restrict it arbitrarily and unnecessarily. Furthermore, the suggested changes threaten to undermine the very purpose for which the Principle was initially formulated and incorporated into legislation, namely to allow regulatory decisions to be made in the face of what are often irreducible uncertainties regarding causal relationships. Decisions made after waiting until uncertainties are reduced through improvements in the knowledge base may be retrospectively preventative, but they are certainly not precautionary.

Gray and Bewers' suggestions (Gray and Bewers, 1996) appear to have been made without reference to the subsequent 1990 and 1995 Declarations on the North Sea. These considerably extended the provisions of the 1987 text. Moreover, no account appears to have been taken of the provisions of OSPAR (1992) which has also significantly influenced the North Sea Ministerial process. In 1990, Ministers reaffirmed their commitment to applying the precautionary principle in the regulation of chemical substances and to the application of policies agreed in 1987 (MINDEC, 1990). They further agreed to reduce inputs of various hazardous substances by between 50 and 70%. The precautionary principle was also applied to the regulation of nutrients, for which regulatory targets were set, and some measures were agreed on the protection of habitats and species and the regulation of fisheries. In 1995, at the Esbjerg meeting (MINDEC, 1995), signatory Governments noted the obligation accepted by all North Sea States to adopt the precautionary and 'polluter pays' principles as a result of signing the 1992 Convention for the Protection of the Marine Environment of the North East Atlantic (OSPAR, 1992) and further strengthened those obligations by committing to a definitive timeline for implementation.

§17. The Ministers AGREE that the objective is to ensure a sustainable, sound and healthy North Sea ecosystem. The guiding principle for achieving this objective is the precautionary principle. This implies the prevention of the pollution of the North Sea by continuously reducing discharges, emissions and losses of hazardous substances thereby moving towards the target of their cessation within one generation (25 years) with the ultimate aim of concentrations in the environment near background values for natural occurring substances and close to zero concentrations for man made synthetic substances.

The 1995 Declaration, therefore, considerably extended the agreements made in 1987 and 1990 in relation to hazardous substances, emphasizing also that, in the definition of hazardous substances as those that are toxic, persistent and liable to bioaccumulate, 'toxicity should be taken to include chronic effects such as carcinogenicity, mutagenicity and teratogenicity and adverse effects on the function of the endocrine system'. The Declaration also recommended that the precautionary principle be applied in fisheries management.

Even at the time of writing, Gray and Bewers' proposals (Gray and Bewers, 1996) were made in relation to a declaration that had been significantly modified by the agreements discussed above that have been made subsequent to 1987. These appear *de facto*

to accommodate the majority of their proposals and to incorporate a precautionary approach to other aspects of marine environmental protection in addition to chemical inputs. None the less, their article raises a number of points concerning the role of science in environmental protection and its relationship to the precautionary principle, which are worth exploring in some greater detail.

Precaution — A Scientific Principle

The inclusion of the precautionary principle within numerous legislative frameworks, including the North Sea Ministerial Declarations, OSPAR Convention, the Rio Declaration 1992, the 1996 Protocol to the London Convention and others, reflects the need within such fora for a mechanism to address the pervasiveness of uncertainty and indeterminacy coupled with the ongoing need to take decisions that facilitate preventative action. It has become a key axiom of international environmental protection initiatives, with a role for science explicitly defined.

By apparently dispensing with the need to establish causality, the principle has been widely criticized as unscientific either in whole or in part. It has been categorized as a wholly political/management philosophy rather than a basis upon which to consider scientific knowledge and uncertainty as part of the overall regulatory process. This view is discussed extensively by Bewers (1995) and also by Gray (1990), Gray *et al.* (1991) and Stebbing (1992). Indeed, Gray (1990) states that the precautionary principle should not be a part of science since, by definition, it does not have to rely on scientific evidence. Dovers and Handmer (1995) suggest that 'the primary use of the Precautionary Principle is to fly the flag of ignorance and uncertainty and hold the policy fort until a framework is developed to improve our overall understanding of the problem, guide the application of available techniques and identify areas where appropriate techniques are lacking'.

From other perspectives, the precautionary principle is seen, more constructively, as a paradigm to resolve some of the tensions inherent in translation of scientific knowledge into policy [see, for example, Wynne and Mayer (1993) and Earll (1992)]. In so far as it represents a means explicitly to recognize fundamental, empirical shortcomings in the science applied to decision-making processes, it can be further argued that the precautionary principle is scientific both in nature and intent (Johnston and Simmonds, 1990; Johnston and Simmonds, 1991). The principle should be viewed not as a substitute for a scientific approach but rather as an overarching principle to guide decision making in the absence of analytical or predictive certainty. In this regard, Gray's view (Gray, 1990) that the principle lies outside of science is misleading. The role of scientific research in the early detection of threats to human health and the environment was explicitly recognized in the core definition of the principle even at its earliest conception (FRG, 1986). In light of this, it is likely that the importance of primary scientific research will increase, rather than decrease, under precautionary legislation.

This intent seems to have been extraordinarily difficult for some scientists active at the science/policy interface to grasp. One charitable analysis is that a precautionary approach to environmental protection has been seen as unscientific (e.g. Gray, 1990) or that the role of science in policy-making is in some way threatened by the adoption of such a principle. Other less charitable views abound. The lack of defined causal relationships is frequently cited in support of the option to defer a regulatory decision until such time as further research becomes available. This results in the sceptical view in some quarters that research is often initiated in place of action to reduce pollution and other impacts, rather than in support of it.

Such deferral places reliance on two basic assumptions:

1. that a greater understanding of the system under study will result, allowing risks to be more clearly defined and quantified; and
2. that the risks arising from precautionary action taken now are greater than the currently undefined risks of taking no action until the results of further investigations become available.

With regard to the first assumption, the opposite is often the case. Research frequently demonstrates that ecosystems possess greater complexities and are harder to define and predict than previously thought. Although much research funded as part of ongoing regulatory processes is undoubtedly valuable, it is rarely able to resolve the 'regulators dilemma' (Bodansky, 1991) which led to its commissioning. The second assumption highlights the fundamental point that a decision to delay action is action in itself and, as such, accepts any consequences that may result.

It would seem, however, that overt resistance to acceptance of the precautionary principle by the scientific community has now been replaced by efforts to reinforce the role of science in policy formulation (Gray *et al.*, 1991; Bewers, 1995) involving, in part, attempts to redefine the precautionary principle (Gray and Bewers, 1996) as noted above. Essentially, what is being proposed is that environmental regulation be predicated upon a risk assessment governed by pessimistic assumptions. Although this may be viewed by some as a mechanism by which the frequency of underprotective decisions may be reduced, it must be recognized that risk assessment captures neither the spirit nor the intentions of the precautionary principle. Within the Esbjerg Declaration, risk assessment is recognized only as a tool to assist in setting priorities and developing action programmes (MINDEC, 1995)

as part of the fundamental commitment to the precautionary elimination of all discharges, emissions and losses of hazardous substances.

Provisions under Annex 2 of the Esbjerg Declaration effectively devolved the development of assessment tools to OSPAR. As part of its work programme, OSPAR is currently developing a strategy in relation to hazardous substances which explicitly defines a role for scientific procedures in the selection, assessment and prioritization of chemical discharges and is developing working definitions of toxicity to take into account further toxicological effects and exposure scenarios. In fact, the risk-based approach currently being discussed within OSPAR conforms very closely with the approach advocated by Gray and Bewers (1996).

Limitations of Risk Assessment

The common framework of risk assessment comprises hazard identification and assessment together with an exposure assessment (Suter, 1993). The output from these two elements is then used to conduct a comprehensive risk assessment. Hazard assessment attempts to define properties inherent to a particular substance or activity in quantitative terms, albeit within the limitations imposed by the range of hazards considered and the methods employed. Risk assessment then turns upon the use of estimates of exposure to the stressor in question in order to arrive at an interpretation of the significance of the identified hazards as threats to the environment or human health.

Clearly, in order for a risk to be managed, it must first be identified and quantified; what cannot be measured cannot, after all, be managed. Notwithstanding the use of a raft of pessimistic assumptions, it must be recognized that the assessment process is immediately constrained by decisions as to precisely which hazards are to be considered. By selection of such criteria, whether through judgmental exclusion or through ignorance of other possible pathways or mechanisms, system uncertainties are artificially reduced (Wynne, 1992). Two important and fundamental questions then arise:

Firstly, have all possible impacts been identified and incorporated into the assessment? Have factors such as the cumulative effects of a particular chemical in conjunction with other stressors (such as overfishing, erosion, etc.) or the additive or synergistic effects of multiple chemical exposures been taken into account? If not, then a proportion of the risks posed by the activity, and an unknown proportion at that, will simply not be recognized or addressed.

Secondly, can all those impacts identified be fully quantified?

The uncertainties surrounding these questions lie at the very heart of the limitations of the risk assessment process.

Often, as noted by Berg and Scheringer (1994), 'it is taken for granted that the relevant events which lead to effects which are then qualified as damage are known', despite the fact that professional judgment plays a significant role in the identification of such events. Objective assessment may, therefore, be compromised by the bias, however unintentional, imposed according to the concerns and perceptions of the assessors. Variations in previous professional experience, coupled with differing judgments as to what constitutes damage, or even an effect, could result in two independent assessments of a single system being substantially different in terms of construction, data requirements and output. Moreover, the interpretation of calculated risks relies on definitions of criteria of acceptability that, in themselves, are generally not subject to the same degree of analysis or assessment (Berg and Scheringer, 1994). Indeed, in many cases, such acceptability criteria are set more or less arbitrarily.

In common practice, risk assessment procedures for the marine environment are highly simplified in form from the ideal. They compare the Predicted No-Effect Concentration (PNEC) with the Predicted Environmental Concentration (PEC) to produce a PEC/PNEC ratio whose magnitude determines subsequent regulatory action. Despite being the accepted formula for risk assessment within, for example, the European Union, generic risk assessments, applicable to clearly defined and described situations and/or localities should be distinguished from a comprehensive risk assessment. In the latter, the concept of risk is expressed as the probability of an adverse event occurring (Bro-Rasmussen *et al.*, 1996) on an ecosystem basis. Such assessments involve a more holistic and, inevitably, a more resource intensive approach [see, for example, Solomon *et al.* (1996)].

In reality, the PEC/PNEC ratio approach differs little from the water-quality criteria approach originally developed by the USEPA (Train, 1979) and that depends upon the use of arbitrarily defined 'safety factors' to establish a 'safe' concentration. Suggestions by Chapman *et al.* (1998) that safety factors should be tuned to the characteristics of dose–response curves, in an effort to improve their scientific defensibility, would undoubtedly introduce a greater degree of complexity and specificity into the assessment. Nevertheless, it would not address the fundamental problems arising from the extension of laboratory toxicity tests, based on a limited number of end-points, to arrive at quantitative predictions of effects and safe concentrations in the field.

The fundamental role of safety factors is, after all, to be protective rather than predictive, in the face of uncertainty or limitations to data. To weight the factors according to one aspect of the hazard assessment for which data availability (subject to the end-points studied and test conditions employed) may be greater, would be to introduce subjective bias to the characteri-

zation of uncertainty that may well compromise their protective role. Increasing the extent to which the dose–response relationships for individual end-points are reflected in setting safety factors may increase the relevance of such factors artificially and should not be seen as a surrogate for the identification and inclusion of other end-points that may have very different dose-responses. Furthermore, the discussion presented by Chapman *et al.* (1998) appears to incorporate the assumption that safety factors will be reduced as the amount of understanding of a system increases. This assumption ignores the possibility that further research may indicate that a hazard had previously been underestimated or simply unidentified.

Despite the simplicity of the safety factor approach, availability of analytical data frequently remains insufficient to support complete assessment. Bro-Rasmussen *et al.* (1996) take the view that the development of risk assessment is impeded by the same lack of data and documentation as impedes the general hazard assessment of chemicals. Given the huge number of individual chemical substances, the development of simplified test methods and the portrayal of generic risk assessment models based on these tests as acceptable does not actually make them robust. Furthermore, Bro-Rasmussen *et al.* (1996) point out that future needs will include better basic investigations, and an intensification of studies designed to generate data for use in hazard and risk assessments. The paucity of data currently available for use in risk assessments is frequently overlooked, resulting in overconfidence in risk assessment. Some of the critical areas are examined below.

Hazard Assessment

Other than the initial problem formulation step that forms part of risk assessment process in the US context, the first substantive step in risk assessment is hazard assessment. The shortcomings of current information resources upon which to base hazard assessment procedures have been strikingly illustrated by a study of chemicals present in the waters of UK estuaries (Matthiessen *et al.*, 1993). Of 71 compounds isolated, some of which were only tentatively identified, Chemical Abstract Services (CAS) identifying numbers could not be found for 29%. Some toxicity data were found for 35% of the compounds isolated and quantitative structure–activity relationships (QSAR) calculations were possible only for a further 33%. Overall notional 'safe' levels using 100-fold application factors could only be estimated, therefore, in the case of 68% of the chemicals, with almost half of these relying on QSAR estimations. The developments in safety factors suggested by Chapman *et al.* (1998) would not appear to provide a resolution to this problem. Even tentative identification of a proportion of chemicals being discharged cannot be routinely achieved. It is quite normal for anything up to 90% of substances isolated by GC-MS screening techniques to remain unidentified (Johnston *et al.*, 1996). Obviously, if it is not possible to identify a particular chemical, then it is not possible to assess its hazardous properties comprehensively. In addition, the isolated compounds are likely to represent only a subset of those actually present due to selectivity of sample extraction procedures and limitations imposed by the specificity of analytical techniques.

In addition, Matthiessen *et al.* (1993) note that the accurate prediction of the joint effects of complex mixtures of substances is not possible at present. This is unfortunate since chemicals entering the environment generally do so as part of a mixture and ultimately join other mixtures of substances already present. Apart from problems that arise as a result of the toxicological interaction of components of the mixture (additivity, potentiation, synergism or antagonism), close covariance of chemical concentrations over an impacted area or variation in the composition of the discharge can also confound attempts to establish causal relationships.

QSARs are increasingly being used where empirical data are not available. However, the major utility of QSARs is as a component of early screening measures and not as a rigid predictor of properties. Chemicals with the ability to disrupt endocrine systems are a case in point. Of the many potential adverse interactions of environmental chemicals with the endocrine system, the oestrogen system has been by far the best studied. As noted by Katzenellenbogen (1995) and McLachlan (1993), the variety of chemical structures that cause oestrogenic responses in sensitive test systems is enormous. There is currently no explanation as to how such a structurally diverse array of chemicals stimulate activity. It has, for example, proven impossible to predict oestrogenicity of the phthalate esters on the basis of their molecular structure alone (Harris *et al.*, 1997).

It must be recognized that although QSAR equations may work well when dealing with structurally similar chemicals within a group, this is not always so. Moreover, when addressing highly heterogeneous classes of toxicants, the equations require a large number of independent variables in order to address all significant molecular features. In consequence, QSARs simply cannot be used to confirm that two chemicals have a similar mode of action and hence to arrive at a ranking of relative hazards, based upon a hazard assessment of only one of them (Vighi and Calamari, 1996). Quite clearly, the utility of QSARs as predictive tools is extremely limited.

The use of the three marker properties of toxicity, persistence and liability to bioaccumulate alone is regarded as inadequate by Gray and Bewers (1996). Chapman (1997) considers it to be a nonsensical combination, noting that persistence and potential to

bioaccumulate are vital properties of certain essential heavy metals and that toxicity is a dose-related phenomenon. Again, this position seems to have been adopted without consideration of the definitions of hazard within regulatory frameworks, such as the North Sea and OSPAR processes, or of the nature of the precautionary commitments under which these are employed. Chapman's concerns (Chapman, 1997) regarding, for example, the inclusion of heavy metals by reference to these properties, are addressed at the implementation stage of these agreements by distinction between the targets required for synthetic and naturally occurring compounds (MINDEC, 1995).

Furthermore, if attempts are made to extend hazard assessments beyond these marker properties, it becomes clear that almost no consistent data are available. Although not ideal, the wide application of toxicity, persistence and bioaccumulative capacity as indicators of hazard stems from a recognition that these are best furnished with data and provide some basis for evaluation of hazard where data on other aspects are largely unavailable.

Exposure Assessment

Despite the undoubted improvements in the detection limits of analytical techniques in recent years, this cannot be regarded as synonymous with their overall adequacy for regulatory purposes. As an example, the observation by Wells (1993) that measurements made on the same sample by one laboratory often bear no resemblance to the values obtained by another laboratory still holds true for many chemicals. Wells (1993) observes that the interpretation of the environmental significance of such data sets may be no more than an observation on the spread of errors generated by the analysts concerned. Only 21 chemicals/chemical groups, not all of them mandatory determinands, are subject to the current QUASIMEME intercalibration programme (Wells, 1993; Wells *et al.*, 1993; Wells and Balls, 1994). This programme was set up to address the concerns associated with analytical accuracy and precision in environmental samples. The number of substances addressed by this initiative is very low when compared with some 50000 substances in use in the EU alone of which 4500 are demonstrably hazardous to aquatic systems (Edwards, 1992). It is estimated that, each year, anywhere between 200 and 1000 new chemicals enter the market place (Shane, 1994).

Even when single determinands are of significance and are the subject of established monitoring regimes, further doubts attach to the capability of many monitoring programmes to the detect trends in environmental concentrations. This is because too few data points are available to give their analysis much statistical power. A good example is the ICES Cooperative Monitoring Programme (CMP) of metallic contaminant levels in fish muscle. Fryer and Nicholson (1993) state that accounting for random between-year variations, it would be unrealistic to expect to detect any patterns of change over a 2–3-yr period. The CMP data sets span no more than 8 yr giving sufficient power to resolve trends in zinc of 10% per year and of copper and mercury of 20% per year assuming that the analytical data can be relied upon. These are large trends. If single incidents are considered, then zinc and copper concentrations would need to change by 200% and mercury by 400% to be identified, given a fixed sample number. Only by extending the CMP data to cover 20 yr does the power of the programme rise to the extent that trends of 5% can be detected with a probability of 90% for most metals, excepting lead and chromium. As pointed out by the authors, the environmental impact of a contaminant over this period prior to a trend emerging could prove unacceptable. The power of surveillance monitoring programmes is not simply a question of analytical precision and accuracy, but is also critically influenced by the frequency and intensity of monitoring.

Gray and Bewers (1996), in addition to other authors (Peterman and M'Gonigle, 1992; Buhl-Mortensen, 1996), have noted the lack of attention paid to statistical power in research and monitoring programmes. Statistical power is defined as the probability of detecting an effect where one exists (Peterman and M'Gonigle, 1992). Lack of attention to power has led to experimental designs that seek to minimize the probability of incorrectly identifying an effect when none exists, i.e. committing a type I error (Underwood, 1997), so as to avoid regulatory regimes that are unnecessarily strict. However, efforts to minimize type I errors lead unavoidably to increases in the probability of type II errors, i.e. not identifying real impacts. A type II error could lead to inadequate legislative protection and failure to prevent adverse impacts on the environment or human health. Experiments that erroneously fail to identify an effect may lead to acceptance of the null hypothesis (no effect), when the experimental design would have lacked sufficient statistical power to have identified an effect in the first place.

Nonetheless determining the statistical power requires the researcher to establish an 'effect size', an estimated magnitude of the effect of concern that could be considered significant. The power of an experiment depends partly upon the effect size determined. Thus, although important, the recommendation for a greater focus on avoiding type II errors (Gray and Bewers, 1996) may not compensate *per se* for problems inherent in scientific method. If the researcher is unable to understand and quantify the extent of impacts caused by environmental contamination, let alone identify which adverse effects to

examine, then a reduction of type II errors will not improve the scientific rigor of the experiment.

Description and Prediction of Impacts

In many cases, prediction of impacts is based upon simple single species toxicity tests. Highly standardized procedures are recognized and widely employed. The familiarity of the format, however, leads to widespread neglect of the inherent weaknesses and assumptions. Essentially, what is being determined using such tests is the distribution of tolerance of a given test species. Data from the tests are generally normalized and used to derive a median effective concentration (e.g. LC_{50}, EC_{50}) to which statistical confidence limits can be attached and the variability of the response can be represented as a linear function. If, therefore, a toxic chemical is present in the environment at sufficient concentration to elicit a detectable response, then not all individuals will be equally affected, all other variables being kept constant.

This concern was highlighted most recently by Evenden and Depledge (1997) in their discussion of variations in genetic susceptibility within ecosystems. These authors point out that even among clones of a single species, sensitivity to chemical exposure and other stresses can vary in both nature and magnitude. Within a clone, some individuals can show greater variations in sensitivity than others. Previous exposure to a contaminant may further enhance genetic variability and susceptibility to this or other stressors during future exposure scenarios.

Moreover, the hazard assessment procedure will only be valid while the dose/response relationship conforms to classical toxicological assumptions. This may not always hold true. In the past decade, the fundamental assumption that magnitude of effect increases with the level of a particular stressor has also been challenged. A test of the impact of contaminated dredge spoils upon haematological parameters of common shrimp (Smith *et al.*, 1995) showed, counter-intuitively, that the greatest effect was exerted by a mixture of 95% clean sand and 5% dredge spoil rather than 100% dredge spoils. In other words, the test showed a marked departure from a monotonic response. A similar phenomenon was also observed by Nelson *et al.* (1993) in invertebrate toxicity tests using clean sand as a diluent for contaminated sediment.

To date, no satisfactory explanation for these observations has been discerned. Nevertheless, they imply that impacts of chemicals associated with sediments could, in some instances, be proportionally greater than expected when they are remobilized and naturally diluted with clean particulate material as may be the case in estuarine systems or around the periphery of dumpsites. Furthermore, since sediments play an important role in the sequestration and remobilization of contaminants, such findings also introduce further uncertainty into the reliability of hazard assessments based upon organisms that live in the water column. Although there is insufficient evidence to suggest that such departures represent the general rule, they introduce further sources of error with obvious implications for the development of tiered test procedures used to evaluate sediment toxicity for regulatory purposes (Adams *et al.*, 1992). Although departures from standard toxicity curves may have been recognized for some time, they do not appear to have gained wide consideration in the development of standardized protocols.

In the case of endocrine disrupting chemicals, the inability of classical dose–response relationships to characterize adequately the behaviour of these chemicals has already been widely recognized. In addition to the other unique challenges posed by the regulation of such chemicals, including the diversity of chemical structures that can elicit responses, they can also show the greatest effects at the lowest doses tested (Patlak, 1996). This can produce an inverted U-shaped curve that cannot be readily subjected to standard statistical analyses. Moreover, the standard regulatory approach of forecasting effects at low doses based on them occurring at high doses becomes invalid. Indeed, the use of a safety factor approach could well result in a chemical being regulated at a level where the maximum biological impact occurs. In addition, an inverted U-shaped dose/response in the case of oestrogenic chemicals may indicate that at high doses, substances are acting as anti-oestrogens or are having effects mediated through other mechanisms. Further, the ability of in-vitro tests to predict whole-organism effects cannot be assumed. Tests for endocrine effects other than oestrogenicity are at a relatively much less advanced stage of development and validation.

As Gray and Bewers (1996) note, considerable progress has been made in elucidating and defining biomarkers of toxicological effect (Depledge and Fossi, 1994) using characteristic responses of individual organisms. Many are based upon developments in molecular biology, but the term biomarker has come to encompass both new methodologies and classical biochemical measures of exposure. This, at least in theory, opens the way to direct evaluation of stress imposed upon organisms in situ.

In practice, the situation is somewhat more complex than this. Firstly, the range of possible indicators of environmental status or change is essentially unlimited. The selection of appropriate markers is a subjective and, in some cases, generally arbitrary process (Berg and Scheringer, 1994) based largely on what is measurable (Wynne, 1992). As with tests based upon simple lethality or more subtle chronic effects, however, the major limitations imposed upon the use of defined biomarkers flow from attempts to translate a relatively simple biochemical change in a single representative organism to prediction of potential impacts upon

complex, whole ecosystems (Weeks, 1995). Figure 1 illustrates this in more detail. Broadly speaking, with increasing levels of biological organization, both in the laboratory and in the natural environment, the depth of understanding achieved tends to lessen. In addition, understanding of natural systems is inevitably less than for tightly controlled experimental systems.

Biomarkers tend to diminish in utility as indicators of causality with the increasing complexity of the system under consideration (Suter, 1994). An organismal or sub-organismal biomarker may well allow an effect to be identified at a population or ecosystem level. The precise cause of the effect, however, may be extremely difficult to resolve from the multiple factors acting at such levels of biological organization (Power and McCarty, 1997). Paradoxically, whereas changes at the sub-organismal or organismal level are the cornerstone of ecological assessments, the protection of populations, communities and ecosystems is of greatest concern to environmental managers. It follows that the choice of both test and assessment endpoints for any given management goal is of critical importance (Suter, 1994).

The responses of the organisms used in toxicity testing are also of considerable importance to the robustness of the hazard and subsequent risk assessment process. Cairns (1986, 1989) outlines two pivotal assumptions made in ecotoxicological studies:

firstly, the assumption that if the most sensitive species are used for testing then all other species will inevitably be protected;

secondly, the assumption that, as a consequence, it is not possible that malfunction at higher levels of organization will occur (Cairns, 1986, 1989).

These assumptions clearly lose their validity entirely if the test organisms differ in their relative sensitivity to different chemicals, or are not actually representative of the most sensitive species. The ability to survive laboratory handling clearly plays a very significant role in the selection of organisms for controlled toxicity testing (Power and McCarty, 1997), and, in turn, it seems likely that tests on the most sensitive species would rarely be feasible, even if that species could be identified. In practice, a relatively limited range of test organisms are routinely employed.

Work by Olsgard and Gray (1996) illustrates the above problems extremely well. Using multivariate statistical analysis, the authors were able to detect subtle, but significant, alterations to the benthic community at distances several kilometers from individual oil drilling installations and over areas of up to 60 km^2 surrounding an oil field. Previous studies based on standard toxicity tests had greatly underestimated the impact of drill cuttings and produced waters on benthic communities. Olsgard and Gray (1996) concluded that this was probably because relatively robust test organisms were used. Therefore, although the laboratory tests provided evidence of dose–response relationships, they were unable to predict community responses accurately. Conversely, whereas multivariate statistical techniques applied to benthic surveys were more capable of detecting subtle spatial changes, they provided only a limited insight into the causal factors involved. Approaches based on the concept of the sediment quality triad (Chapman *et al.*, 1997) recognize the value of parallel collection of chemical, biological and toxicological data from a number of stations in the construction of a desriptive matrix of sediment characteristics over a wide area. Such approaches may allow more detailed correlative analysis of relationships between variables, but cannot overcome the difficulties in establishing causality, particularly when many chemical and biological factors covary.

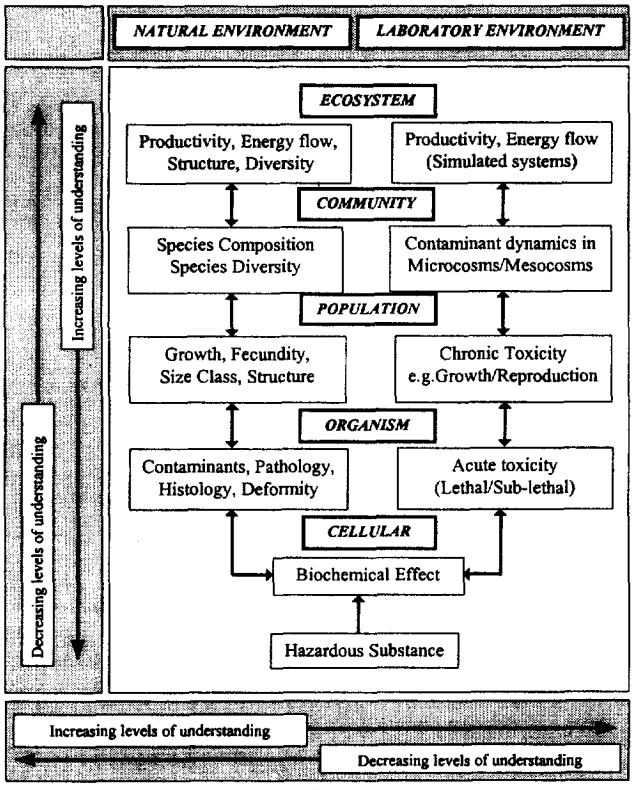

Fig. 1 Diagramatic representation of the relationship between level of biological organization and degree of system understanding for both laboratory and natural environments. In general terms, as the level of organization increases and control of conditions decreases, so our understanding of system functions and interactions declines.

Implications for Risk Assessment

The above general deficiencies strike at the heart of all risk-assessment frameworks. The inherent properties of a substance cannot be determined if it cannot be identified. If it can be identified, the hazard assessment procedure is likely to be held hostage to the adequacy of existing informational resources. The determination

of probable exposure in the environment is contingent upon the existence of adequate analytical techniques as well as knowledge of the inherent properties of the substance and its behaviour in the environment after discharge. The complexity of effluents and the consequences for chemical interactions, fates and toxicological activity (Johnston *et al.*, 1996) are rarely considered. In common practice, therefore, hazard tends to be assessed in terms of simple toxicological effects and simplified fate scenarios. LC_{50} values are most frequently determined in single species toxicity tests. Persistence is measured in terms of an environmental half-life. The potential to bioaccumulate is determined either from direct measurement or by use of a surrogate such as the log octanol/water partition coefficient. This approach is exemplified by the UK approach to prioritizing chemicals for regulation in the 'Red List' (DoE, 1988) where weighting is also given to the overall production volume of the chemical concerned. The EU and current OSPAR approaches are broadly similar. In the case of the EC, progress in evaluating chemicals has been painfully slow (Edwards, 1992). From the original EC priority candidate list of 129 chemicals, only 18 had been regulated by daughter directive of the original Directive of 1976 (76/464/EEC) by 1992. The prioritization methodology also effectively means that chemicals are ultimately being regulated on the somewhat dubious basis of extrapolating the results of single species toxicity test results to the wider environment.

Power and McCarty (1997) summarize the range of common misconceptions that have become established in ecological risk assessment, including the two assumptions outlined above regarding the selection of 'the most sensitive species' and the problems of extrapolation of test data on single species to higher levels of biological organization. They note that these will contribute individually or in combination to significant errors in gathering, analyzing and interpreting data used to describe environmental problems. In their opinion, 'regulatory decision-making successes based on the extrapolation of laboratory results should not be seen as confirmation of the scientific validity of extrapolation, but rather as a victory for common sense and sound judgment.' They point out that the interpretation of relevance of toxicity tests requires an insight into the functioning of the whole ecosystem in terms of multiple anthropogenic and natural stressors and the interactions between those stressors but that ecologists do not currently understand which factors are the most critical.

Perhaps one of the most interesting of the ancillary misconceptions listed by Power and McCarty (1997) is that 'good science' will remedy any and all environmental problems. Clearly, there is an ongoing need for pure and applied knowledge to facilitate improved understanding of ecosystem function and contaminant behaviour (Ducrotoy and Elliott, 1997), but such knowledge will always be incomplete and, in itself, can form only part of responsible policy and management systems.

Risk assessment essentially treats all uncertainties as if they were due simply to incomplete definition or description of the system (Wynne, 1992) and does not recognize the existence of indeterminacies, i.e. unknowns, such as mechanism of action or causal chains, that are simply not amenable to analytical reduction. Indeed, as Wynne (1992) emphasizes, indeteminacies are not simply large uncertainties. Rather, they are unidentified but inherent components within the definitions of risks or uncertainties and on which these quantifiable measures are unavoidably conditional. Berg and Scheringer (1994) point out that ecosystems possess underlying dynamism and trends that cannot be distilled down to simple relationships and examined to reveal their detailed complexity. Similarly, Power and McCarty (1997) highlight the 'derivative properties' of an ecosystem that are not discernible from its component parts. Moreover, it is not possible to define fixed boundaries between ecosystems, nor natural reference conditions against which damage can be compared, other than by subjective judgment. Predictions of ecosystem function and potential impact may be confounded not only by inaccuracies in measurements or estimates of the assessment criteria under study, but also by factors acting on the system from outside its artificially defined boundaries. In addition, baseline studies rarely give an understanding of the historical development and underlying dynamics of a system. These characteristics, encapsulated by Berg and Scheringer (1994) in the term 'overcomplexity', are inherent properties of natural systems.

Gray and Bewers' approach (Gray and Bewers, 1996) assumes that the uncertainties discussed in the preceding sections will be compensated for by the use of pessimistic assumptions. These will, inevitably, be subjectively derived and likely to be based upon professional judgment albeit supported by some basic data on the chemical concerned. They are, therefore, likely to be no more reliable than optimistic assumptions, or indeed capable of being differentiated from them. After all, neither body of assumptions can take into account the complexity of the receiving system, deficiencies in the methods available to estimate the hazard or indeed lack of knowledge of the actual existence of a hazard. Indeed, in its assumption that all factors of significance have been addressed, risk assessment is, in itself, a highly optimistic approach. In practice, pessimistic assumptions are indeed applied in certain areas; the safety factor approach used to extrapolate from animal toxicity tests to derive tolerable daily intakes (TDIs) by humans of hazardous chemicals or as part of the water quality criteria approach (Train, 1979) are perhaps the best examples, although underestimation of hazard may result in these

later proving lower margins of safety than initially assumed. This is not a precautionary approach, but rather one based in management terms on the assumption that safe margins of exposure can be defined with reference to effect concentrations in laboratory studies.

This then is the background to the definitional changes proposed by Gray and Bewers (1996). They are based on procedural and methodological precepts intended to reduce reliance upon what they regard as poorly substantiated perceptions of effect. For the most part, these precepts have demonstrable shortcomings in their power to predict or detect environmental change at the higher levels of biological organization. Given this, there is a real danger that their use in the regulatory decision-making process could lead to 'precautionary action', as defined by Gray and Bewers (1996), becoming hostage to the availability of scientific data. Notwithstanding adoption of pessimistic assumptions and a concomitant recognition of the existence of uncertainties, this could result in the burden of proof of deleterious effect continuing to fall upon those seeking to establish a rationale for regulatory action in response to chemical-, site- or region-specific concerns. Similarly, action based solely upon the availability of scientific evidence could lead to decisions being based, as they have been in the past, upon the flawed assumption that absence of evidence of an effect is analogous to evidence of absence of impact.

However well intentioned the changes proposed by Gray and Bewers (1996) actually are, they require substantial departure from a precautionary approach. This may lead, in turn, to misconception of the purpose underlying precautionary action and, moreover, may result in misuse of the paradigm by those wishing to sustain a 'business as usual' approach.

Conclusion: The Need for Precaution

Power and McCarty (1997) suggest that society has been forced to use science to understand anthropogenic influences upon the environment and that the tool of risk assessment has been developed in response to these needs and resource limitations. They further note that, although risk assessments are widely used, consensus on an acceptable and comprehensive framework that clearly identifies the roles of policy and science in decision-making has not emerged. Moreover, the question of who bears responsibility for a decision based on an assessment of risk subsequently found to be incorrect remains to be answered (Funtowicz and Ravetz, 1994). Indeed, as Gray and Bewers (1996) point out, those who make incorrect predictions are rarely penalized. The impasse reached has led to the surfacing of numerous tensions. Many ecologists regard ecotoxicology as an overly simplistic representation of reality and hence fundamentally flawed. At the other end of the spectrum, some environmental groups are contemptuous of science as currently applied at the policy/regulatory interface, seeing it in the pejorative terms outlined above. The two views are commonly grounded.

Jackson and Taylor (1992) stress that policy-makers are almost always required to act in the presence of some considerable uncertainty or doubt. In this regard, it must be recognized that, although scientific investigation and interpretation are information providers, uncertainties and indeterminacies will always remain. Scientific analyses and deductions may serve policy-makers with valuable information that can be used to identify hazards and to prioritize and guide decisions (Funtowicz and Ravetz, 1994), but cannot replace the decision-making process itself (Wynne, 1992; Power and McCarty, 1997).

As noted by Berg and Scheringer (1994), 'acting in an environmentally sustainable way is more than just applying ecological knowledge'. A risk deemed acceptable according to ecotoxicological assessment and defined environmental acceptability criteria may, nevertheless, be unacceptable in political or social terms. As Bewers (1995) notes, acceptability needs to be judged on grounds other than applied science. The size of the decision stakes and the potential consequences of being wrong are also essential considerations when arriving at policy decisions (Wynne, 1992).

Clearly, there is a need to develop regulatory cultures that encourage wider and more involved debate of the relative benefits, costs and uncertainties of policies and decisions, particularly those relating to ecosystem protection and risks to human health. This does not imply simply that output from risk assessments (e.g. for chemical exposure) should be interpreted for public consumption by comparison with risks associated with day-to-day activities. The fact that these exercises involve comparison of risks estimated with large uncertainties, with those determined empirically from existing statistics, is conveniently ignored. Invariably, such comparisons are highly subjective, although few attempts are made to communicate to the public this subjectivity or the uncertainties surrounding the risk assessment. Rather, it is essential that inherent indeterminacies and differences in values and perceptions are recognized and that transparent mechanisms exist to address them. Furthermore, although pure and applied sciences have an essential role to play in the identification of issues, the description of systems and impacts and the ranging of hazards, they alone cannot drive policy.

The basis and necessity of an alternative approach, particularly in relation to global environmental concerns, is extensively discussed by Funtowicz and Ravetz (1994), under the term 'post-normal science'. Inherent to this approach is the involvement of an 'extended peer community' in the evaluation not only of the quality of scientific data but also the application of these data, and associated value judgments, in the

formulation of policy. A similar view is expressed by Ducrotoy and Elliott (1997), who recognize the need not only for scientific information to be effectively disseminated but also for involvement of scientists, politicians, NGOs and the wider public in reaching final decisions. As the authors point out, such an approach may be essential to facilitate 'the management of irreducible uncertainties in knowledge and ethics' (Funtowicz and Ravetz, 1994).

The precautionary principle has been developed to reconcile the widening gap between the theories of ecotoxicology and the pragmatic needs of the regulators who bear the responsibility to act as far as possible to avoid damage before it has occurred. In issues of environmental contamination, chemical exposure and potential impacts on wildlife and human health, it is frequently the case that the facts are uncertain, values in dispute and the stakes high (Funtowicz and Ravetz, 1994). With respect to the need to avoid harm in advance, decisions may also be urgent, particularly as the point at which environmental damage becomes irreversible is practically impossible to predict. To assume that risk assessment, albeit with pessimistic assumptions, can replace the role of precautionary action is to make unreasonable and unrealistic demands on standard scientific methodology. The utility of a risk-based approach in prioritizing issues for regulatory action has been recognized (MINDEC, 1995). However, precaution is an essential characteristic of any responsible decision-making process, particularly when systems are complex, uncertainties are high and the potential consequences of underestimating, or simply not identifying, risks are severe. The precautionary principle, although essentially guided by input from analytical and predictive science, nevertheless recognizes and compensates for inherent uncertainty and indeterminacy in natural systems and provides a central paradigm for responsible, timely and definitive preventive action. The precautionary principle cannot and should not be subsumed under a risk assessment mechanism. Neither should risk assessment be seen as a means of implementing the precautionary principle.

The precautionary principle is, in its own right, a crucial scientific tool to mitigate environmental threats. Its inclusion as the guiding principle within numerous international conventions for the protection of the environment has been in recognition of the need for a mechanism to address uncertainties and limitations to scientific knowledge at the science–policy interface. Redefinition of the principle along the lines suggested by Gray and Bewers (1996) would fundamentally change its function but would not eliminate the need for a precautionary mechanism. Ultimately, such a change would simply necessitate the replacement of the principle with a new paradigm, under a different name, to perform essentially the same function.

Adams, W. J., Kimerle, R. A. and Barnett, J. W. (1992) Sediment quality and aquatic life assessment. *Environmental Science and Technology* **26**, 1865–1875.

Berg, M. and Scheringer, M. (1994) Problems in environmental risk assessment and the need for proxy measures. *Fresenius Environmental Bulletin* **3**, 487–492.

Bewers, J. M. (1995) The declining influence of science on marine environmental policy. *Chemistry and Ecology* **10**, 9–23.

Bodansky, D. (1991) Scientific uncertainty and the precautionary principle. *Environment* **33**, 4–5, 43–44.

Bro-Rasmussen, F., Lokke, H., Kristensen, P., Laursen, E., Boyd, H. E., Jorgensen, C. E., Nielsen, K. M. (1996) *The Non-assessed Chemicals in the EU. Report and Recommendations from an Interdisciplinary Group of Danish Experts*. Danish Board of Technology 1996/5.

Buhl-Mortensen, L. (1996) Type II statistical errors in environmental science and the precautionary principle. *Marine Pollution Bulletin* **32**(7), 528–531.

Cairns, J. (1986) The myth of the most sensitive species. *BioScience* **36**, 670–672.

Cairns, J. (1989) Applied Ecotoxicology and Methodology In *Aquatic Ecotoxicology: Fundamental Concepts and Methodologies*, Volume 2, eds A. Boudou and F. Ribeyre, pp. 275–290. CRC Press, Boca Raton, FL.

Chapman, P. M. (1997) The precautionary principle and ecological quality standards/objectives. *Marine Pollution Bulletin* **34**, 227–228.

Chapman, P. M., Anderson, B., Carr, S., Engles, V., Green, R., Hameed, J., Harmon, M., Haverland, P., Hyland, J., Ingersoll, C., Long, E., Rodgers, J. Jr., Salazar, M., Sibley, P. K., Smith, P. J., Swartz, R. C., Thompson, B. and Windom, H. (1997) General guidelines for using the Sediment Quality Triad. *Marine Pollution Bulletin* **34**(6), 368–372.

Chapman, P. M., Fairbrother, A. and Brown, D. (1998) A critical evaluation of safety (uncertainty) factors for ecological risk assessment. *Environmental Toxicology and Chemistry* **17**(1), 99–108.

Depledge, M. H. and Fossi, M. C. (1994) The role of biomarkers in environmental assessment. *Ecotoxicology* **3**, 161–172.

DoE (1988) *Inputs of Dangerous Substances to Water: Proposals for a Unified System of Control. The Government's Consultative Proposals for Tighter Controls over the Most Dangerous Substances Entering the Aquatic Environment (The 'Red List')*. Publ. UK Department of the Environment, July 1988, 25 pp+Annexes.

Dovers, S. R. and Handmer, J. W. (1995) Ignorance, the precautionary principle and sustainability. *Ambio* **24**(2), 92–97.

Ducrotoy, J.-P. and Elliott, M. (1997) Interrelations between science and policy-making: the North Sea example. *Marine Pollution Bulletin* **34**, 686–701.

Earll, R. C. (1992) Commonsense and the precautionary principle — an environmentalists perspective. *Marine Pollution Bulletin* **24**(4), 182–186.

Edwards, P. (1992) *Dangerous Substances in Water, A Practical Guide*. Environmental Data Services, London 64 pp.

Evenden, A. J. and Depledge, M. H. (1997) Genetic susceptibility in ecosystems: the challenge for ecotoxicology. *Environmental Health Perspectives* **105**(Supplement 4), 849–854.

Fryer, R. J. and Nicholson, M. D. (1993) The power of a contaminant monitoring programme to detect linear trends and incidents. *ICES Journal of Marine Science* **50**, 161–168.

FRG (1986) *Umweltpolitik: Guidelines on Anticipatory Environmental Protection*. Federal Ministry for the Environment, Nature Conservation and Nuclear Safety, 43 pp.

Funtowicz, S. O. and Ravetz, J. R. (1994) Uncertainty, complexity and post-normal science. *Environmental Toxicology and Chemistry* **13**(12), 1881–1885.

Gray, J. S. (1990) Statistics and the Precautionary Principle. *Marine Pollution Bulletin* **21**, 174–176.

Gray, J. S., Calamari, D., Duce, R., Portmann, J. E., Wells, P. G. and Windom, H. L. (1991) Scientifically based strategies for marine environmental protection and management. *Marine Pollution Bulletin* **22**, 432–440.

Gray, J. S. and Bewers, J. M. (1996) Towards a scientific definition of the precautionary principle. *Marine Pollution Bulletin* **32**(11), 768–771.

Harris, C. A., Henttu, P., Parker, M. G. and Sumpter, J. P. (1997) The estrogenic activity of phthalate esters *in vitro*. *Environmental Health Perspectives* **105**, 802–811.

Jackson, T. and Taylor, P. J. (1992) The precautionary principle and the prevention of marine pollution. *Chemistry and Ecology* **7**, 123–134.

Johnston, P. and Simmonds, M. (1990) Precautionary Principle (Letter). *Marine Pollution Bulletin* **21**, 402.

Johnston, P. and Simmonds, M. (1991) Green light for precautionary science. *New Scientist*, 3 August, 4.

Johnston, P. A., Stringer, R. L. and Santillo, D. (1996) Effluent complexity and ecotoxicology: regulating the variable within varied ecosystems. *Toxicology and Ecotoxicology News* **3**, 115–120.

Katzenellenbogen, J. A. (1995) The structural pervasiveness estrogenic activity. *Environmental Health Perspectives* **103**(Supplement 7), 99–101.

McLachlan, J. A. (1993) Functional toxicology: a new approach to detect biologically active xenobiotics. *Environmental Health Perspectives* **101**, 386–387.

Matthiessen, P., Thain, J. E., Law, R. J. and Fileman, T. W. (1993) Attempts to assess the environmental hazard posed by complex mixtures of organic chemicals in UK estuaries. *Marine Pollution Bulletin* **26**, 90–95.

MINDEC (1987) *Ministerial Declaration of the Second International Conference on the Protection of the North Sea*. 24–25 November 1987, London, UK.

MINDEC (1990) *Ministerial Declaration of the Third International Conference on the Protection of the North Sea* 7–8 March 1990, The Hague, The Netherlands.

MINDEC (1995) *Ministerial Declaration of the Fourth International Conference on the Protection of the North Sea*, 8–9 June 1995 Esbjerg, Denmark.

Nelson, M. K., Landrum, P. F., Burton, G. A. Jr., Klaine, S. J., Crecelius, E. A., Byl, T. D., Gossiaux, D. C., Tsymbal, V. N., Cleveland, L., Ingersoll, C. G. and Sasson-Brickson, G. (1993) Toxicity of contaminated sediments in dilution series with control sediments. *Chemosphere* **27**, 1789–1812.

Olsgard, F. and Gray, J. S. (1996) A comprehensive analysis of the effects of offshore oil and gas exploration and production on the benthic communities of the Norwegian continental shelf. *Marine Ecology Progress Series* **122**, 277–306.

OSPAR (1992) Final Declaration of the Ministerial Meetings of the Oslo and Paris Commissions. Oslo and Paris Conventions for the Prevention of Marine Pollution, Paris, 21–22 September 1992.

Patlak, M. (1996) A testing deadline for endocrine disruptors: EPA scrambles to develop a screening program for these complex substances. *Environmental Science and Technology* **30**, 540A–544A.

Gonigle, 1992" Peterman, R. M. and M'Gonigle, M. M. (1992) Precautionary principle 2. Statistical power analysis and the precautionary principle. *Marine Pollution Bulletin* **24**(5), 231–234.

Power, M. and McCarty, L. S. (1997) Fallacies in ecological risk assessment practices. *Environmental Science and Technology* **31**, 370A–375A.

Shane, B. S. (1994) Introduction to Ecotoxicology. In *Basic Environmental Toxicology*, eds L. G. Cockerham and B. S. Shane, pp. 3–10. CRC Press, Boca Raton, FL.

Smith, V. J., Swindlehurst, R. J., Johnston, P. A. and Vethaak, A. D. (1995) Disturbance of host defence in the common shrimp *Crangon crangon* by exposure to harbour dredge spoils. *Aquatic Toxicology* **32**, 43–58.

Solomon, K. R., Baker, D. B., Richards, R. P., Dixon, K. R., Klaine, S. J., La Point, T. W., Kendall, R. J., Weisskopf, C. P., Giddings, J. M., Giesy, J. P., Hall, L. W. and Williams, W. M. (1996) Ecological risk assessment of atrazine in North American waters. *Environmental Toxicology and Chemistry* **15**, 31–76.

Stebbing, A. R. D. (1992) Environmental capacity and the precautionary principle. *Marine Pollution Bulletin* **24**, 287–295.

Suter, G. W. (1993) Ecological Risk Assessment. Lewis Publishers, Boca Raton, FL.

Suter, G. W. (1994) Endpoints of interest at different levels of biological organisation. In *Ecological Toxicity Testing: Scale Complexity and Relevance*, eds J. Cairns and B. R. Niederlehner, pp. 35–48. Lewis Publishers, Boca Raton, FL.

Train, R. E. (1979) *Quality Criteria for Water*, Castlehouse Publications, Tunbridge Wells, UK, 256 pp.

Underwood, A. J. (1997) Environmental decision-making and the precautionary principle: what does this principle mean in environmental sampling practice?. *Landscape and Urban Planning* **37**, 137–146.

Vighi, M. and Calamari, D. (1996) Quality objectives for aquatic life: the problem of mixtures of chemical substances. *Human and Ecological Risk Assessment* **2**, 412–418.

Weeks, J. M. (1995) The value of biomarkers for ecological risk assessment: academic toys or legislative tools?. *Applied Soil Ecology* **2**, 215–216.

Wells, D. E. (1993) QUASIMEME — An Introduction. *QUASIMEME Bulletin* No. 1 March 1993. Publ. SOAFDS, Aberdeen, UK.

Wells, D. E., Cofino, W. P., Quevauviller, P. and Griepink, B. (1993) Quality assurance of information in marine monitoring: a holistic approach. *Marine Pollution Bulletin* **26**, 368–375.

Wells, D. E. and Balls, H. R. (1994) [Editors] QUASIMEME: quality measurements for marine monitoring. Progress on the first year of the EC Project (1993). *Marine Pollution Bulletin* **29**, 143–249.

Wynne, B. (1992) Uncertainty and environmental learning: reconceiving science and policy in the preventative paradigm. *Global Environmental Change*, June 1992, 111-127.

Wynne, B. and Mayer, S. (1993) How science fails the environment. *New Scientist*, 5 June, 33–35.