CHAPTER 17

Toxicological versus Ecotoxicological Testing

GUIDO PERSOONE AND JAMES GILLETT

17.1 ANALOGIES AND DIFFERENCES BETWEEN TOXICOLOGY AND ECOTOXICOLOGY

Evolution has gradually brought one species, namely man, above other living creatures on Earth (at least as far as intelligence is concerned). It is logical and justified, therefore, that concerns about the negative effects of man-made chemicals on humans still rank much higher than ecotoxicological considerations in hazard assessment. In contrast to toxicology, which focuses on the protection of the human species (considered by themselves to be the most important creatures on Earth), ecotoxicology concerns the protection and well-being of several millions of species scattered over a variety of terrestrial and aquatic habitats, and of the biological communities and their encompassing ecosystems and processes. Thus, ecotoxicology also indirectly addresses the well-being and ultimately the survival of mankind since noxious chemicals can degrade ecosystems to the extent that the environment may no longer fulfil basic human needs.

We recognize the universality of living processes since we recognize that all life forms have certain basic features in common. However, we are faced with a very wide variety of types and forms of life. Experience has shown that, in general, the greater the differences in form and function, the greater the differences in response to foreign chemicals. These differences are the basis for selective toxicity and are brought about by differences in (i) the mode of uptake of chemicals, (ii) rates of chemical penetration, metabolism, and excretion, (iii) chemical reaction at the toxic site and the mechanism of that reaction, and (iv) the specific habitats of the biota that determine the potential for exposure to chemicals. These same features are important in human toxicology in determining the risk to specific segments of society but the range of responses is obviously much greater in ecotoxicology.

Human toxicological research is, by definition, restricted to the species level and most of our interest is directed at the individual. However, much information is extrapolated from research with animals as well as human epidemiology data. Ecotoxicology, by contrast, encompasses the species and infra-species levels, as well as the impacts on the structure and function of biota at supra-organismal levels.
(i.e. populations, communities and ecosystems), all of which may be subject to an infinite variety of environmental variables and interactions.

Whereas the loss of individual human lives is of great concern, the disappearance of a few individuals in a population of plants or animals is usually not considered to be serious as long as the proper functioning of the population, communities or ecosystems remains unaffected. Ecotoxicology, therefore, focuses attention on processes and interrelationships of structure and function in populations, communities and ecosystems.

Cairns (1980) emphasized that although natural ecosystems may have so much functional redundancy that the disappearance of certain species does not necessarily lead to impairment of function, it is possible to impair ecosystem function without actually killing organisms.

17.2 EXPOSURE ASSESSMENT VERSUS EFFECTS (HAZARD) ASSESSMENT

Determination of the ‘hazard’ which a particular chemical may represent for man and other biota is usually approached by relating the magnitude (concentration), frequency and duration of exposure to the magnitude, extent and duration of toxic effects which the xenobiotics have on the living creatures.

Exposure, on the other hand, is determined by a variety of factors including the distribution or partitioning of the chemical between environmental media (such as soil, sediment, water and air), the transformation or degradation of the chemical, and the movement of media containing these agents. Ascertaining these factors is ‘exposure assessment’ which dictates the magnitude of the potential hazardous effects.

Once the hazard assessment has qualitatively identified the nature of adverse effects and a quantitative dose–response relationship has been defined, it is possible to assess the risk to components, processes and individuals within the system. The most important feature of risk assessment, as emphasized in all recent reviews on the subject, is the determination (or estimation) of the extent to which concentrations of a given chemical released into the environment (i.e. exposure) overlap in time and space with those that are toxic (i.e. hazardous) to selected organisms, populations and ecosystems. Currently, both exposure and the effects of xenobiotics in the natural environment can only be estimated approximately and in a very crude way because of the infinite number of variables involved. The variables arise from the complex biological structure of ecosystems and the numerous interrelationships between their components, and the multivariant pathways of distribution of chemicals in the environment.

17.3 TERRESTRIAL VERSUS AQUATIC ENVIRONMENTS

Ecologically, the terrestrial environments, in many aspects, are similar to aquatic environments since the living components of both are structured in the same way.
Plant and animal species are grouped in populations which, in turn, are well-organized into biological communities. Both types of ecosystems fulfill the same basic functions; namely, primary (photosynthetic) production, consumption and growth as secondary productivity, and degradation and nutrient cycling as major links throughout several types of food chains.

Even though a considerable part of the species which populate terrestrial and aquatic ecosystems are very closely related phylogenetically, the approach to studying impacts of chemicals is, nevertheless, very different for these two environments. This results from the quite different modes of exposure of biota to chemicals in the two types of ecosystems.

In terrestrial ecosystems, exposure to pollutants present in the air seems only to directly affect the plant kingdom; animals are indeed mainly contaminated by uptake of toxicants via the food chain. In the aquatic ecosystem, the pelagic fauna and flora are in continuous contact with chemicals dissolved or suspended in the water column; food chain uptake appears to be very secondary or at least much slower than that from direct contamination. Suspended solids in aquatic environments have about as much contribution to aquatic species as does dust to terrestrial species. In both types of ecosystems, however, the biota living in the soil/sediments can be intoxicated by contact exposure and ingestion of contaminated particles. This more intense contact provides mobility to these deposited chemicals as they enter food webs and disturbed materials re-equilibrate with the water column.

The quantities of chemical substances that can be carried by each of the media (air, water or soil) are not comparable. Moreover, their dynamics, bioavailability and ultimate fate in each medium can be radically different. As a result, experimental approaches to determine the potential effects of pollutants are quite different for terrestrial and aquatic organisms and will be treated separately elsewhere in this volume.

17.4 THE NECESSITY FOR AND DIFFICULTY OF TESTING AT DIFFERENT LEVELS OF BIOLOGICAL ORGANIZATION

In 1983, Cairns addressed the question, ‘Are single species toxicity tests alone adequate for estimating environmental hazard?’ He noted that ‘at each succeeding level of biological organization, new properties appear that would not have been evident even by the most intense and careful examination of lower levels of organization’. Cairns correctly concluded that, to date, ‘no scientifically justifiable evidence exists to indicate that degree of reliability with which one may use single species tests to predict responses to higher levels of biological organization’.

Until a few years ago, bioassays on non-human target organisms were restricted almost exclusively to laboratory tests on selected test species (of even age or size) exposed for fixed periods of time to constant concentrations of a toxicant in a predetermined set of physiochemical conditions. The major reason for this is that ecotoxicology, historically, started as an offspring of human toxicology. Originally,
the interest was in toxicity ranking of chemicals for a very limited number of test species which were of direct economic or aesthetic interest to man. Standardization of assay conditions provided support for litigation in the control of aquatic pollution (Mount and Gillett, 1982) and in pesticide regulation (Tucker and Crabtree, 1970). Interestingly, these tests had little in common with the reality of impacts on populations and communities which receive stochastic exposure to multiple toxicants under varying physiochemical conditions without reference to species, age or size, or other factors.

The growing concern about pollution in the 1960s, fortunately paralleled by a rapid development of ecology, made mankind aware of the need for analysis of the potential impact of chemicals on natural environments at the ecosystem level rather than at the single species level. All the interacting biological components of the systems at risk are exposed, either simultaneously or successively, to varying concentrations of xenobiotics for different periods of time under site-specific, but also variable, environmental conditions. Thus, both direct and indirect effects were expected and confirmed.

Extensive research during the last decade has shown that the ‘ecosystem approach’ to hazard assessment is very difficult to apply in practice. The difficulty arises from the number of variables involved in multispecies testing either in the laboratory (microcosm or mesocosm) or under actual environmental conditions (field studies). This is a major obstacle to the repeatability of the tests and, as a result, to their predictive potential.

From the practical point of view, there is an undeniable inverse relationship between ‘ecological realism’ on one hand, and simplicity of testing on the other, as illustrated in Figure 17.1. This inverse relationship sustains the dictum that an understanding of toxic mechanisms is achieved by progressing downward in complexity (i.e. from the organism to the physiologic level, to the organ and finally to the molecular level), whereas the outcome of the toxicologic event can only be fully understood by progressing upward to the more complex system. Enzyme inhibition is only of significance if it causes organ and physiologic systems to fail in ways that make the organism less capable of functioning within the community. The allegory, ‘for want of a nail . . . the kingdom was lost’, illustrates the vital concept of a critical pathway for a constituent of a system to result in an adverse effect; not every lost nail results in kingdoms falling.

This concept is particularly relevant in the context of this review. It is possible to focus on mechanisms in a single organism while understanding the functional relationships of many component systems of higher biological complexity. For example, a set of separate in vitro tests may establish that the chemical penetrates the organism, survives in its various systems, and inhibits a key enzyme. If we know that the enzyme is found in specific organs with functional connections to adverse outcomes, we do not need to test each individual organism to predict that adverse effect. In the human species, an acetylcholinesterase inhibitor effectively deprives the heart of enervation, depriving the brain of oxygen and resulting in death. In other
species, the same enzyme system may have different mechanistic connections, the organs may be more or less redundant and important to organismic vitality, and the individual may have only an infinitesimal role in the population, community or ecosystem. Thus, our testing at lower levels helps us to understand the selective nature of toxicity, the comparative aspects of physiology, and perhaps even something about functional relationships within populations and higher systems. However, the same kind of testing tells us little about the actual outcome; our knowledge of the ecological relationships is so incomplete that any conclusions drawn in this respect may be little more than speculation.

A decade or more ago, compensation for ignorance was achieved by use of safety factors or application factors (Mount and Gillett, 1982). Pragmatically, standards set using such arbitrary factors as 100 or 1000 were effective in reducing pollution and restoring biologic ‘livability’ to the environment. We do not yet know whether this approach was economically invalid (over-regulation) or ecologically inappropriate (e.g. undiscovered loss of species due to inappropriate standards).

During the decades of the 1960s and 1970s, species- and habitat-specific effects were discovered which revolutionized thinking about environmental problems. Direct toxicity of pesticides to song birds and fish was among the first environmental watersheds crossed, but that was soon overshadowed by the chronic lethal toxicity
to robins feeding on contaminated earthworms, which in turn was surpassed by concern about non-lethal, chronic toxicity of pesticides and other toxic substances to birds and fish as seen in reproductive and behavioral effects. These served as warnings to man of potential dangers to his own well-being and illustrated ecological connections previously ignored in anthropocentric haste.

A crude but effective multispecies system was devised by Metcalfe et al. (1971)—a ‘farm pond’ model ecosystem or microcosm—to demonstrate these connections. Although this particular test system has been labelled an ‘ecological junk heap’, had it been employed before the chlorinated hydrocarbon insecticides were introduced, DDT would never have been used on such a spatial and mass scale over most of the globe. At about the same time, various other systems were excised from fields and ponds to examine, among other things, ecosystem response to stress and radioactive fallout. Together, these sets of tests stimulated multispecies testing and microcosm technology as a legitimate approach to evaluation of higher biologic functions and outcomes.

17.5 SINGLE SPECIES TESTS

Although experience with multispecies testing is being gained rapidly, standardized protocols for multispecies tests are only just becoming available. Because reliable, well-tested and/or cost-effective experimental ecosystem level tests which can be used on a routine basis have been lacking, ecotoxicologic testing to support all national and international hazard assessments of chemicals is still based entirely on single species tests, mostly in a tiered approach.

In principle, the candidate species should be representative of the various biota of natural ecosystems, have a comparable sensitivity, possess the functional community connections, and be sufficiently comprehensive or inclusive to be representative of all critical parts of natural ecosystems. Unfortunately, we are far from achieving this goal. Only a few species have been or can be maintained in the laboratory. Buikema and Benfield (1979) point out that the lack of ecological information needed to establish laboratory populations is a serious handicap to increasing the variety of potential candidate species for ecotoxicological testing. Animal species most widely used for testing in the USA include five mammals, five birds, seven fish, two terrestrial insects, nine aquatic arthropods and one mollusc (Kenaga, 1978).

Most tests so far have been carried out on fish, not only because this group is considered to be the best understood for the aquatic environment, but also (and perhaps mainly) because of its direct interest to man (Cairns, 1982). Also, it appears that many of the species routinely used for toxicity testing are, among their fellows in the natural environment, those with the lowest ecological needs and the largest tolerances regarding many environmental variables (euryspecies); this implicitly makes them the easiest to maintain and handle in the laboratory.

It should be emphasized that the methodology and standardization of testing
procedures, even for those of a routine nature, is often very poor. Maki (1983) indicates that although few intercalibration exercises have been carried out, most have been a complete failure either because the experimental protocols have been inadequate or because of the failure to understand some of the basic ecological requirements of the test species. For the aquatic environment, only three tests have successfully passed a 'round robin' inter-laboratory comparison; these are the

Figure 17.2 Interrelationships of the basic factors determining the choice of bioassay test methods (from Persoone, 1980).
**Daphnia magna** and the *Brachydania rerio* short-term bioassays for the freshwater environment (Cabridenc, 1979a,b) and the *Artemia salina* short-term test for the marine environment (Vanhaecke and Persoone, 1982).

The 'ecological realism–simplicity' inverse relationship for the ecosystem approach is, unfortunately, also valid at the single species level. Persoone (1980) clearly demonstrated the paradox between the biological and the economical approach in selecting species and setting up protocols for toxicity tests (Figure 17.2). The biological approach (which is the logical one, of course) aims at the maximum ecological representativeness and sensitivity of the tests. The selection of representative test species (difficult to maintain and handle) and sensitive criteria (growth or reproduction) automatically leads to complicated test methods and long-term experiments which are expensive.

The economical approach, which is the only realistic one for routine testing of a large number of chemicals, is based on simple, short-term tests with limited numbers of species which are easy to culture and on criteria which are simple to assess (e.g. lethality). As noted earlier, this testing strategy, while having merit in terms of economy and expediency, is lacking in terms of representativeness, completeness and sensitivity.

Macek *et al.* (1978), who reviewed the types of aquatic toxicity tests that provide useful data for estimating or predicting the toxicity of chemicals to aquatic organisms, worked out a very informative matrix based on six evaluation criteria (ecological significance, scientific and legal defensibility, availability of routine methods, predictive utility, general applicability and simplicity, and costs); this is reproduced in Table 17.1. Lethality, despite its apparent crudeness, is the most useful of all endpoints considered, both in a single species as well as multispecies tests.

### 17.6 NECESSITY OF EXPOSURE EVALUATION

In order to make the best use of single species tests in predicting how a particular chemical may influence biota under natural environmental conditions, the inherent toxicity of the particular xenobiotics (determined experimentally) should be related to the quantity of the chemical that will reach the organism under the prevailing conditions of exposure. In recent hazard assessment projects (e.g. OECD, 1982), exposure is determined as 'predicted environmental distribution' (PED) and 'predicted environmental concentration' (PEC). The PED is based on a number of physicochemical characteristics (such as water solubility, vapour pressure, relative molecular mass, soil adsorption coefficient and octanol–water partition coefficient). This gives an estimate of the distribution of the chemical within the environmental compartments of air, water, soil/sediment, and biota. The complementary PEC aims at quantitative prediction of the real concentrations of the chemical likely to be achieved within distinct parts of the ecosystem.

The PED is based on a limited number of parameters. It is usually simple to determine. The PEC is much more difficult to estimate since it requires information
Table 17.1  Consensus evaluation of the relative utility of simple and complex toxicity tests used in assessing the risk associated with the occurrence of chemicals in aquatic environments (from Macek et al., 1978)

<table>
<thead>
<tr>
<th>Test</th>
<th>Ecological significance</th>
<th>Scientific and legal defensibility</th>
<th>Availability or routine methods</th>
<th>Predictive utility</th>
<th>General applicability</th>
<th>Simplicity and cost</th>
<th>Present relative utility</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute lethality</td>
<td>33</td>
<td>36</td>
<td>35</td>
<td>22</td>
<td>24</td>
<td>26</td>
<td>100</td>
</tr>
<tr>
<td>Embryo/larval</td>
<td>33</td>
<td>30</td>
<td>23</td>
<td>20</td>
<td>25</td>
<td>14</td>
<td>82</td>
</tr>
<tr>
<td>Reproduction</td>
<td>36</td>
<td>30</td>
<td>21</td>
<td>27</td>
<td>24</td>
<td>8</td>
<td>82</td>
</tr>
<tr>
<td>Residue accumulation</td>
<td>27</td>
<td>32</td>
<td>22</td>
<td>24</td>
<td>17</td>
<td>14</td>
<td>77</td>
</tr>
<tr>
<td>Algal assay</td>
<td>30</td>
<td>23</td>
<td>26</td>
<td>18</td>
<td>14</td>
<td>20</td>
<td>74</td>
</tr>
<tr>
<td>Organoleptic</td>
<td>14</td>
<td>25</td>
<td>25</td>
<td>27</td>
<td>13</td>
<td>15</td>
<td>67</td>
</tr>
<tr>
<td>Structure/activity</td>
<td>18</td>
<td>21</td>
<td>23</td>
<td>17</td>
<td>16</td>
<td>22</td>
<td>66</td>
</tr>
<tr>
<td>Behavioural</td>
<td>21</td>
<td>13</td>
<td>10</td>
<td>9</td>
<td>16</td>
<td>9</td>
<td>44</td>
</tr>
<tr>
<td>Histological</td>
<td>10</td>
<td>10</td>
<td>20</td>
<td>10</td>
<td>13</td>
<td>10</td>
<td>41</td>
</tr>
<tr>
<td>Physiological and biochemical</td>
<td>12</td>
<td>10</td>
<td>13</td>
<td>8</td>
<td>19</td>
<td>8</td>
<td>40</td>
</tr>
<tr>
<td>In vitro</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>4</td>
<td>11</td>
<td>11</td>
<td>25</td>
</tr>
<tr>
<td><strong>Simple systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field</td>
<td>34</td>
<td>24</td>
<td>18</td>
<td>21</td>
<td>21</td>
<td>3</td>
<td>69</td>
</tr>
<tr>
<td>Diversity</td>
<td>26</td>
<td>19</td>
<td>25</td>
<td>15</td>
<td>20</td>
<td>10</td>
<td>65</td>
</tr>
<tr>
<td>Benthic</td>
<td>21</td>
<td>18</td>
<td>12</td>
<td>18</td>
<td>17</td>
<td>12</td>
<td>56</td>
</tr>
<tr>
<td>Microcosm</td>
<td>19</td>
<td>10</td>
<td>15</td>
<td>14</td>
<td>17</td>
<td>9</td>
<td>48</td>
</tr>
<tr>
<td><strong>Complex systems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
on volumes of production, discharge patterns and detailed knowledge of receiving waters (flows, characteristics, locations, etc.).

17.7 ‘SHORT-TERM’ ECOTOXICOLOGIC ASSAYS

We desire to protect species which will never be tested because they are endangered (e.g. ospreys), physically difficult if not impossible to culture (blue whales, redwood trees), or simply not known. Such an organism might be a ‘keystone species’ on which ecosystem function depends (e.g. a detritivore that facilitates nutrient cycling in a pond or forest, or a top-level predator structuring the ecosystem by its prey selection). Even to identify such an organism may require years of field and laboratory study. Rarely do toxicity studies provide an opportunity to evaluate impacts on these organisms and the systems they support, because to do so leads to experiments in ‘biogeochemical time’.

The short-term view of man in ecological and biogeochemical studies is at the heart of ecotoxicology. Failure to take a long view of the outcomes of anthropogenic inputs into the environment results in problems being discovered after they have progressed to near eco-catastrophe. Global pollution by acid precipitation and stratospheric modification are often cited as examples of such phenomena for which both preventative and mitigative responses are quite difficult. Synthetic organic chemicals, such as DDT, PCBs and phthalate esters, are now widely distributed in the environment. These chemicals have accumulated and biomagnified in some ecosystems, and may gradually change the species composition of an ecosystem. Strenuous efforts have restored species of fish or birds to some such habitats, but others still suffer from the impacts of chemicals long after they have been ‘regulated’ or ‘banned’.

Hence, ecotoxicologic assays not only require multispecies forms and interactions, but they also require multiseasonal studies of succession, adaptation and other interactions. In theory, one can select species with short reproduction and life cycles, such as Daphnia magna L. (generation time of days) or Arabidopsis spp. (seed-to-seed time of about 30 days). However, that simply begs the question ‘Are the critical species affected long-lived or do they have multi-seasonal requirements for critical functions?’. The period during which multi-species test systems and microcosms can be maintained in a suitable operational state is often far short of desirable. How then might ‘short-term’ assays for ecotoxicologic effects be developed? There are few courses from which to choose; ignore long-term (multiseasonal or multi-generational) phenomena; ignore multi-species interactions; ignore cryptic species and processes; or assume that broad safety factors (100 or 1000) are adequate. Pragmatically, as noted earlier, the application of safety factors has served well for a time, but as economic issues intruded into environmental protection in the wake of a global energy crisis and slump in both industrial and developing nations, the question of over-regulation is not a rhetorical one.

Unpublished studies by Mount and his co-workers, and similar analyses by others
Toxicological versus Ecotoxicological Testing

using large databases, have revealed some features of comparative toxicity that provide cautious encouragement for the possibilities of ecotoxicologic testing within a shorter timeframe. For example, it is possible to test ‘clusters’ of organisms and extrapolate to a value of a theoretical LCO (the lethality to the zeroth organism) (Gillett, unpublished). This ‘cluster’ is a set of species (e.g. rainbow trout, D. magna, and bluegill, or white rat, bobwhite quail, and housefly) for which commensurate doses (mg/kg) or exposures (mg/m³) are available and for which there is some understanding of the relationship of toxic response to the class of chemical involved in the test. Thus, for heavy metals, a sensitive micro-organism (such as E.coli) could be included. If phytotoxicity is expected or a concern, an alga might be used.

The spread of LD50 or LC50 data from the cluster is used to estimate selective toxicity, while the mean value estimates its qualitative nature (i.e. very highly toxic, moderately toxic, etc.). Provisional analysis of a large database from the US Department of the Interior, Fish and Wildlife Service reveals that there is greater variation between, say, rainbow trout under a variety of test conditions (temperature, age or size, hardness, salinity, pH, etc.) than there is among all fish tested (Johnson and Finley, 1980). Thus, one suspects that any laboratory test is not likely to reveal specific sensitivities, but simply ranks chemicals among species qualitatively. About as much information is gathered from the test cluster (designed to provide taxonomic spread in response) as from more extended testing. However, when selectivity is high (wide range in response) or a pollutant is suspected to affect several media, then the use of a larger cluster (5–8 species) is advisable.

A second approach assumes that chronic lethality occurs at about an order to magnitude lower in exposure than acute toxicity. Non-lethal but serious effects on reproduction and behaviour may be two orders of magnitude lower than lethality or morbidity. Without a sound understanding of chemical ecology and the behavioural relationships, one can never be sure that disruption will not occur as a direct or indirect result of the chemical in the environment. However, in practice, the use of this estimated chronic value (1/10 of the lowest LC50) does protect many species. However, we are never confident about how many species are protected or of their functional importance.

Multi-species and microcosm studies provide a short-cut to investigating these functional relationships, even though these tests frequently require weeks to months for completion and analysis. Once preliminary single species tests and physiochemical data have provided estimates of chronic exposure and chronic response, the model ecosystem can be used to confirm the activity of a candidate chemical or mixture or to demonstrate its probable safety, at least under the conditions of the test. We do not know with what certainty one can extrapolate from a given test set (species, habitat conditions) to others. Therefore, systems most likely to be exposed and/or anticipated to be sensitive to a particular class of chemical should be tested first.

Attention must be paid both to broad functions (primary production, nutrient
cycling, etc.) and species composition. For example, Harte et al. (1980) found that their lotic microcosm tracked primary production and nitrogen cycling in the source reservoir — even beyond the time when a sensitive diatom was eliminated (due to a deficiency in silicon) and replaced by a blue-green alga. The presence of the undesirable organism (Anabaena sp.) would not have been found by simple functional analysis.

This review began by emphasizing that the objectives in ecotoxicology differ markedly from those in human toxicology. We have attempted to demonstrate concerns about approaches to testing which depend on biological simplification for quick answers. At the same time, the complexity of the systems and the species which form them demand greater attention that can only be given to them if we can find economical, accurate and simple means of determining their responses to concentrations of potential toxicants.

REFERENCES


