BELLE NEWSLETTER

Biological Effects of Low Level Exposures

A Publication for the Northeast Regional Environmental Public Health Center, University of Massachusetts, School of Public Health, Amherst, MA 01003
Vol. 14 No. 3, January 2008, ISSN 1092-4736

HORMESIS AND ETHICS

Introduction

Edward J. Calabrese
Department of Public Health
Environmental Health Sciences
University of Massachusetts
Amherst, MA 01003, USA
Phone: 413-545-3164; Fax: 413-545-4692
E-mail: edwardc@schoolph.umass.edu

There are numerous possible ethical issues that emerge in the process of risk assessment and risk communication. These issues could include the use of animal models in the hazard assessment process, the protection of high risk groups, the use of unverified assumptions on the risk assessment process and many more. Recognition of ethical questions and their careful assessment is an important consideration for society and the scientific community. This is not an inherently easy or comfortable topic as there can be significantly different views by equally well meaning people on critical topics. We see this in numerous areas of societal life, with risk assessment being no exception. Even though there is little likelihood for broad based and unifying agreements on many ethical issues, there is a need for these issues to be recognized, discussed and refined. It was within this context that a decision was made to have an entire issue of the BELLE Newsletter devoted to the topic of Hormesis and Ethics.

Hormesis is a biological concept about the dose response. It describes a dose response relationship in which there is a low dose stimulation and a high dose inhibition. It is believed to be highly generalizable, being independent of biological model, chemical class and endpoint measured (Calabrese and Baldwin, 2001; Calabrese, 2005a). A "problem" with hormesis is that it challenges dose response models that have been used as the mainstays for environmental risk assessment for noncarcinogens and carcinogens (Calabrese and Baldwin, 2003; Calabrese, 2004; Cook and Calabrese, 2006a). Such challenges have not gone unnoticed as a variety of papers have offered criticisms of the hormesis concept, especially with respect to its application to risk assessment (Axelrod, et al., 2004; Thayer

TABLE OF CONTENTS

et al, 2005; see rebuttals - Calabrese, 2005b; Cook and Calabrese, 2006b). Some of the risk assessment questions posed related to ethical concerns and hence underlie the importance of the current attempt to examine hormesis within a ethics framework.

To accomplish this goal Dr. Kevin Elliott, University of South Carolina, was invited to help create and then edit this issue of the BELLE Newsletter. Dr. Elliott accepted the responsibility of identifying experts in the area of ethics and environmental risk who might like to apply their thinking to the concept of hormesis. With the exception of the Hoffmann and Stempsey paper which was invited by the BELLE office, all authors were invited by Dr. Elliott. Dr. Elliott developed a series of questions (see below) that were to either be addressed directly by the authors or to guide the direction of their intellectual inquiries. After receiving all invited author contributions Dr. Elliott wrote an integrative summary which is the last article of this issue. As the readers will see, this is an intriguing area that is just beginning to be explored and it is our hope to return to this general topic in the future. All manuscripts were published without editing by BELLE.

Edward J. Calabrese

QUESTIONS PROPOSED TO EXPERTS:

by **Dr. Kevin C. Elliott,** *University of South Carolina*

(1) When considering the possible alteration of regulatory policy in response to hormesis, what do you regard as some of the most important ethical considerations? Taking those considerations into account, are there any changes to public policy that you would or would not recommend based on the information that we currently have about hormesis?

(2) What ethical principles are most appropriate for guiding public policy in the face of scientific uncertainty, and how might those principles apply to the hormesis case? For example, would some formulation of the precautionary principle be ethically advisable in this case or not, and what would its ramifications be?

(3) What ethical considerations should guide public policies that have to balance potentially positive health impacts of toxins on some individuals against potentially negative health impacts of the same toxins on other individuals? What about balancing positive effects of a toxin on some biological endpoints against negative effects on other endpoints, or positive effects on some time scales against negative effects over other time scales?

(4) Are there particular sorts of future scientific studies that you would recommend, based on their potential to provide important empirical information that would facilitate more ethically informed decisions about applying hormesis to public policy?

(5) From an ethical perspective, would you recommend integrating something like "public participation" or "broad-based deliberation" (see the NAS volume Understanding Risk) into discussions about hormesis and regulatory policy? If deliberation is warranted, what are some promising mechanisms and guidelines for implementing it in this case? Should scientific researchers who study hormesis be aware of any particularly important ethical responsibilities related to facilitating this sort of deliberation or providing information to deliberative bodies?

REFERENCES

Axelrod, D., Burns, K., Davis, D., and Von Larebeke, N. (2004). Hormesis – An inappropriate extrapolation from the specific to the universal. *Int. J. Occup. Env. Hlth.*, 10:335-339.

Calabrese, E.J. (2005a). Paradigm lost, paradigm found: The re-emergence of hormesis as a fundamental dose response model in the toxicological sciences. *Env. Poll.*, 138:378-411.

Calabrese, E.J. (2005b). Hormesis – Basic, generalizable, central to toxicology and a method to improve the risk assessment process. *Int. J. Occup. Env. Health*, 10:476-477.

Calabrese, E.J. (2004). Hormesis: from marginalization to mainstream. A case for hormesis as the default dose-response model in risk assessment. *Toxicol. Appl. Pharmacol.*, 197:125-136.

Calabrese, E.J. and Baldwin, L.A. (2003). Toxicology rethinks its central belief – Hormesis demands a reappraisal of the way risks are assessed. *Nature*, 421(6924):691-692.

Calabrese, E.J., and Baldwin, L.A. (2001). Hormesis: U-shaped dose-response and their centrality in toxicology. *Trends in Pharmacol. Sciences*, 22(6):285-291.

Cook, R.R., and Calabrese, E.J. (2006a). The importance of hormesis to public health. *Env. Hlth. Perspect.*, 114:1631-1635.

Cook, R.R., and Calabrese, E.J. (2006b). Hormesis is biology, not religion. *Env. Health Perspect.*, 114:A688-A688.

Thayer, K.A., Melnick, R., Burns, K., Davis, D., and Huff, J. (2005). Fundamental flaws of hormesis for public health decisions. *Env. Hlth. Perspect.*, 113:1271-1276.

SCIENCE, REGULATION, AND HORMESIS

Heather Douglas
University of Tennessee
808 McClung Tower
Knoxville, TN 37996-0480, USA
Phone: 865-974-7206; Fax: 865-974-3509
E-mail: hdouglas@utk.edu

INTRODUCTION

Hormesis is back. After languishing in the backwaters of scientific research for over fifty years, the idea that small amounts of toxic substances can be beneficial to organisms is once again discussed prominently on the pages of scientific journals.[1] A pattern is being seen across multiple chemicals and across multiple species, plant and animal. Some chemicals, at very low doses, seem to help living cells defend themselves. What should this result mean for the regulation of chemicals?

First, we must be clear about the extent of benefit hormesis can provide. It can be misleading to describe the hormetic response as a U-shaped dose response curve. An examination of the curves does not reveal a U-shaped curve, but rather a J-shaped curve. The J-shaped curve is crucial in showing that at low doses, the stimulation of response is much milder than on the right-hand (high dose) side of the curve. Describing the dose-response curve as U-shaped can mislead one into thinking that the benefits available from a hormetic response are equivalent to the harms we find at the higher doses. Let us be clear from the outset: we are not talking about benefits at the low dose that are equal to the harms at the higher doses. Indeed, we are talking about a mild stimulation of the *same* response as found at higher doses. There is no symmetrical U here. Instead, we see an asymmetrical J.

Such a J-shaped response curve has been seen across a wide variety of contexts. That some substances elicit some kind of hormetic response seems clear. What is less clear is what this should mean for the scientific research agenda and for the regulation of chemicals in the environment. It is on these issues that I will focus.

CHANGING THE RESEARCH AGENDA

The idea that hormesis is a fundamental and widespread biological response has received both serious attention and critique. Among the substantive critiques of hormesis, it has been questioned whether the overviews of the pervasiveness of hormetic responses are properly capturing the implications of the original data, and whether the meta-analytic "testing" done leads to statistical problems with the results.[2]

But let us suppose that the results culled from the existing toxicology literature are robust: That in 40% of the studies which could possibly show a hormetic effect (i.e. studies with adequate low dose testing), a hormetic effect is found, i.e. that a J-shaped dose-response curve is recorded.[3]

If this is correct, it is a powerfully suggestive result. It has led hormetic advocates to make some strong claims on its behalf. For example, Karl Rozman has argued that "the accumulated evidence has now reached a critical mass sufficient to postulate that at low doses all chemicals have hormetic... effects..., although in many instances such effects may be immeasurably small."[4] He goes on in the same paper to claim that, for research in toxicology, "the only logical and rational way to go about it is to accept the aforementioned dictum that all chemicals have hormetic effects at low doses."[4] Yet to accept such a dictum in science is far from rational. Indeed, if we accept the dictum, and view the absence of evidence for it in some cases as merely an indictment of our ability to measure the hormetic effects (because they are "immeasurably small"), we have effectively created an unfalsifiable hypothesis, the antithesis of rationality in science.[5]

Perhaps instead of accepting hormesis as a dictum, we should consider accepting hormesis as a default assumption, particularly in light of its apparent explanatory power. There are deep ethical problems with accepting hormesis as a default assumption in risk assessment for regulatory purposes, as I will discuss in the next section. For now, let us consider the possibility of accepting it as a default assumption for scientific research purposes only.

The ability of hormesis to explain data from a broad array of biological processes seems to be good reason for scientists to accept the hormesis thesis.[3,6] Scientists, however, should be wary of the seductive nature of this type of inference. Although the scope and explanatory power of the theory are attractive, such attributes in themselves are no reason to think the theory true. The history of science is littered with theories of broad explanatory power that have had to be substantially modified, trimmed back, or wholly rejected as a more complex picture of the world emerged. While explanatory power and scope are traditional values in scientific research, they should be indicators of research projects that will be fruitful, i.e. allow for investigations into new and unknown territories, rather than indicators of hypotheses that are likely to withstand the test of time.

Consider, for example, the 17th century hypothesis that all physical processes could be understood through mechanical interactions, such as one ball hitting another, or springs, or rods tying objects together. This mechanical approach to scientific research was extremely fruitful in the 17th and early 18th century, helping scientists to organize their experience, to look for new data, and to think through the scientific problems they faced. By the late 18th century, however, the universality of mechanical explanations had been rejected. Newton's gravitational theory was not actually mechanical, positing the then mysterious "force" of gravity, which could act at a distance. with no obvious mechanical basis serving as the intermediary. Chemistry based on mechanical interactions had to be abandoned as well, as mechanical thinking failed to provide adequate resources for understanding the complexity of chemical interaction. What was once thought a univer-

sal principle, because of its scope, explanatory power, and fruitfulness, was scaled back to an approach to a particular area of physics (e.g. the ideal gas laws).

Similar examples can be found in the history of science, from the adherence to wave theories of light and the search for ether in the 19th century to the overly simplistic biological essentialism of early hormone theories in the 20th century. Even if one has a theory that seems to explain a broad range of similar phenomena, one should not infer from that fact that it is true. Rather, such a theory should be thought of as suggestive for which questions one should pursue next, being always open to the fact that future data collected may serve as the basis for refinement, or limitations on, or even wholesale rejection of that very theory.

There are two areas in particular I would suggest scientists interested in the hormesis thesis pursue. The first is to attempt to predict when we can expect hormetic responses to occur. We have not yet seen such responses universally. So scientists interested in hormesis should attempt to make predictions about when and where we should see hormesis in action, predictions that if borne out, will bolster our confidence in the reliability of the hormesis thesis.

The second is to develop accounts of when we can expect hormetic responses to be beneficial to the organism. Scientists working on hormesis have pointed to examples where hormesis (i.e. the J-shaped curve) is present, but where the stimulation response at low doses is not beneficial to the organism.[7,8,9] Thus even if we assumed that hormesis is a universal biological response, given current evidence, we should not assume that the stimulatory response of hormesis is always beneficial.

In sum, we need a hormesis thesis that can predict both when it will occur (and when it will not), and when it will be beneficial (and when it will not). As many others have suggested, understanding the biochemical mechanisms by which hormesis will act will be central to developing the theory further.[2,7,10,11] Given its explanatory power and scope, the theory should be pursued by scientists. However, it should not be taken as predictively successful (the key issue for regulation) until it has better developed its predictive capacities.

CHANGING THE REGULATORY AGENDA

I have suggested that scientists should pursue the hormesis thesis in their research with gusto. It may indeed help develop a much deeper understanding of the biochemistry of life. In the meantime, how should the regulatory community approach hormesis?

Seeing a J-shaped pattern across a range of toxicology tests is suggestive for regulators. Indeed, if one believes that the best available evidence should inform the dose-response curve, which should then inform regulatory standards, it seems that hormesis ought to be taken very seriously in the regulatory arena. Although some advocates of hormesis have argued that hormesis should serve as a default assumption for setting regulatory standards, this argument overstates the case for hormesis. Even the best overviews of currently available data at sufficiently low doses show a hormetic-type response only half the time. And, as noted above, there is significant doubt that all hormetic-like responses are beneficial.

Despite these empirical concerns, it could be argued that the hormesis thesis has as much evidence for it as the two main alternatives: the linear dose-response model and the threshold model. We could attempt to do a serious weighing of which among these models is the best supported from all relevant sources of evidence: toxicological, epidemiological, and biochemical. Indeed, part of the difficulty we have is that all three models have plausible biochemical underpinnings in at least some cases. Rather than tackle this daunting task, I will instead turn to ethical problems that hamper any use of the hormesis thesis in regulation, even if it is the best supported evidentially. These problems are largely avoided by the linear model and completely avoided by the threshold model, which may be why regulators prefer those models.

The ethical problems are best approached by examining two different answers to the difficult question: what purpose is served by the regulation of human exposure to chemicals? There are two general answers to this question: 1) that regulation is to help produce maximum health benefit to the public at minimum cost; and 2) that regulation is to help prevent one party from harming another without their knowledge or consent, i.e. to prevent market failures. It is important to see that these two goals are not equivalent. In the latter, we are attempting to prevent harm to people, imposed for the convenience and benefit of others. Any substantial harms imposed, without foreknowledge and consent, would be unacceptable. In the former, we attempt to weigh the harms imposed on some for the benefit of others to see if the benefits outweigh the harms. If the benefits do outweigh the harms, then the imposition of harms on some for the benefit of others is acceptable.

Under these two different views of what regulation is for, the desirability of taking into account the potential benefits of hormesis looks very different. Indeed, the first view is open to some important concerns about the regulatory process, concerns that illuminate why the burden of proof has fallen to the advocates of the hormesis thesis as a default assumption in risk assessment.

To attempt to take into account the potential benefits of hormesis through a regulatory approach is to gamble with the health of the winners and losers who will be affected by the regulatory standards, particularly standards which attempt to make use of the beneficial hormetic window. The benefits to some will impose costs on others, and it is unlikely either will know who is who in the final result, making compensation or voluntary avoidance impossible. This raises moral concerns about whether it is right to benefit some by imposing harms on others, particularly when no one at increased risk can be warned to take extra precautions.

That attempting to maximize benefit through applying hormesis requires such trade-offs becomes clear when one considers the complications of assessing the risk of exposure across a human population, particularly the complications that arise from the variability among individuals' sensitivity and the variability in individual exposure to

other potentially relevant chemicals (i.e. mixtures). The most sensitive may be harmed at the ideal hormetic dose for the population as a whole. Those with exposure to other substances that act along the same biochemical pathway will be subject to potentially harmful overexposure. Thus, to attempt to regulate to optimize exposure across the population will create winners and losers.

In addition, in order to weigh the benefits of the winners against the harms to the losers requires the use of some dubious economic valuations. We must decide how to compare the health benefits to some against the health harms to others. While economists have developed sophisticated ways to measure the average monetary valuation of life and health across a population, individuals vary in their acceptance of such valuations. If one individual's health is harmed, and they value it far more than the result of an average willingness-to-pay study, why should they be forced to accept this harm in the name of a public good from which they do not benefit, with inadequate or no compensation for that harm?

Finally, the attempt to maximize benefit across a population usually ignores issues of justice. Those most sensitive to chemical exposures are usually the young, the already weakened, the poor, and those with least access to good medical care. The same groups are also often those who have elevated exposure levels to relevant mixtures. These ethical problems, involving the ideas that we should not trade the health of one person for the health of another and that involuntarily vulnerable populations should be protected as a matter of justice, create difficulties for the idea that we should regulate chemicals to maximize public health.

The alternative regulatory purpose of preventing one party from harming another avoids many of these difficulties. If one is solely preventing harm, one does not need to weigh the potential health benefits for some against the potential harms to others. Thorny issues of justice recede. In addition, one does not need to decide how much a life, even a statistical life, is worth, or how much avoiding an illness is worth. These problematic areas become peripheral, and the focus instead is on what level of exposure will avoid the imposition of serious harm, to the best of our knowledge. The potential benefits are irrelevant to the regulators. If specific people want the potential benefits of chemical exposure, they can seek such exposures voluntarily (perhaps through marketed supplements).

Under the more limited purpose of regulation, the hormesis model collapses to the threshold model (assuming the hormetic effects are beneficial). The regulators are solely concerned with which doses cause harm, and regulating to avoid that harm. The threshold model is very useful for such purposes, clearly delineating acceptable from unacceptable doses. If the low dose effects are not beneficial, a different extrapolation method will be needed. In these cases, regulators will have to decide upon an acceptable level of de minimus risk (e.g. 1 in a million lifetime risk). Such a de minimus level of risk does raise some ethical difficulties (why is this level of risk acceptable?), but it is far less fraught than the morass created by the attempts to maximize benefit over an entire population.

One might note that not all of our regulatory efforts with respect to chemicals fit the more limited agenda. The fluoridation of water, for example, is often justified by a desire to produce maximal good for the public. Yet the baseline public resistance to this beneficence should not be discounted as merely irrational fears of being fed a rat poison. The public tends to want to look after their own health goods for themselves, and to want the government to do no more than keep involuntary harms away. The more rational strains of the fluoride debate emphasize concerns over sensitive subpopulations to fluoridation, and the insistence that if parents want this benefit for their children, they should expose their kids to fluoride themselves. That the argument in this case generally turns on the need for society to protect children from potentially negligent parents (as it does in the schooling debates), indicates that this is not a strong counterexample to the more limited regulatory approach. Support for this more limited regulatory approach can also be seen in the debates over smoking in public, where concern over harms to non-smokers, especially workers, are the primary winning argument to ban smoking on airplanes, in hotels, and in restaurants. The smokers can smoke themselves to death, as long as they don't force harm on the person next to them in the process.

The more limited goal of regulation also frames such key laws as the Clean Air Act, which demands *not* that the air produce maximal public benefit, but rather that it prevent harm, within a margin of safety.[12] Similarly, the Safe Drinking Water Act requires efforts to produce water that is as close as is feasible to producing no known harm. The act that regulates food additives also attempts to prevent any involuntarily imposed harms.[13] Why are these acts framed this way? Because breathing air, drinking municipal water, and eating purchased food are not seen as areas where we have much ability to prevent involuntary exposures. The regulatory framework in these cases seeks to protect the public from harm only, not to maximize benefit for the public as a whole.

CONCLUSION

This discussion has been focused on the health effects of chemical exposure and the moral shape of the regulatory framework for that issue. I don't think that the regulatory framework for this issue is readily expandable to other policy issues, such as land use policy, education, or transportation. In those realms, the pursuit of a public good (as opposed to the avoidance of public harms) does seem to drive the structuring of policy and the arguments that the public finds persuasive. However, I would caution against an importation of such ideals into the regulatory health arena. There are good reasons against such an importation of these ideals, including the injustice of imposing health effects on some for the benefit of others, and problems with the economic methods for valuing health and life. People want more information about their food, nutrition, and medical options, so that they can make these decisions for themselves. The value of autonomy over our bodies and health is rarely trumped by the pursuit of a public good.

This creates a high burden of proof for hormesis advocates seeking the use of hormesis as a default assumption in risk assessments. The

potential injustices created by seeking an optimized exposure, injustices to the more sensitive or to the overexposed via mixtures, would have to be ameliorated by at least a more thorough understanding of when and why hormesis occurs. If scientists pursue the biochemical research agenda to unpack the mechanisms and the sources of variability in human populations, perhaps someday we could have a world in which individuals could assess their personal exposure levels, their own genetic susceptibility levels, and could determine their ideal exposure levels to various toxins. Indeed, we could have chemical companies marketing small doses of toxins as health supplements. But the regulatory agenda may always be limited by the purposes for which we seek regulatory action.

REFERENCES

1 Calabrese EF. Historical blunders: how toxicology got the dose-response relationship half right. *Cell Mol Biol* 2005; **51**: 643-654

2 Kitchin KT, Drane JW. A critique of the use of hormesis in risk assessment. *Hum Exp Toxicol* 2005; **24**: 249-253.

3 Calabrese EF, Baldwin LA. Hormesis: the dose-response revolution. *Annual Rev Pharmacol Toxicol* 2003; **43**:175-197.

4 Rozman KK. Hormesis and risk assessment. *Hum Exp Toxicol* 2005; **24**: 255-257.

5 Popper K. Science: Conjectures and Refutations. *Conjectures and Refutations*. London and New York: Routledge, 2002: 43-78.

6 Calabrese EF, Baldwin LA. Hormesis: U-shaped dose responses and their centrality in toxicology. *Trends Pharamacol Sci* 2001; **22** (6): 285-291.

7 Pickrell JA, Oehme FW. Examining the risk and benefits of replacing traditional dose-response with hormesis. *Hum Exp Toxicol* 2005; **24**: 259-264.

8 Pickrell JA, Oehme FW. Examining the risks and benefits of considering both the traditional dose-response and hormesis in arriving at an acceptable exposure level. *Hum Exp Toxicol* 2006; **25**:23-27.

9 Weltje L, vom Saal FS, Oehlmann J. Reproductive stimulation by low doses of xenoestrogens contrasts with the view of hormesis as an adaptive response. *Hum Exp Toxicol* 2005; **24**: 431-437.

10 Thayer KA, Melnick R, Burns K, Davis D, Huff J. Fundamental flaws of hormesis for public health decisions. *Environ Health Perspect* 2005; **113** (10): 1271-1276.

11 Kaiser J. Sipping from a poisoned chalice. *Science* 2003; **302**: 376-379.

12 Fiorino D. *Making Environmental Policy*. Berkeley, CA: University of California Press, 1995.

13 National Research Council. *Risk Assessment in the Federal Government: Managing the Process*. Washington DC: National Academy Press, 1983.

RESPONSE TO QUESTIONS ON HORMESIS AND ETHICS

Sven Ove Hansson
Dept. of Philosophy and History of Technology
KTH
Teknikringen 78
100 44 Stockholm, Sweden
Email: soh@kth.se

1. When considering the possible alteration of regulatory policy in response to hormesis, what are the most important ethical considerations? Based on those considerations, are there any policy changes that arguably should or should not be made in the near future?

Two major ethical problems present themselves in the context of hormesis: How should we weigh negative effects against positive effects? And how should we act when there is uncertainty about the effects?

The weighing of negative against positive effects is commonly described as weighing risks against benefits. Risks are uncertain negative effects, whereas by benefits we usually mean certain positive effects. Therefore, the risk-benefit terminology is somewhat misleading in cases when there is more uncertainty concerning the positive than the negative effect. At the present level of knowledge, this seems to apply in many cases of potential hormesis effects.

From a regulatory point of view, hormesis gives rise to the challenge of regulating a substance that has predominantly positive effects at low dose levels and predominantly negative effects at higher levels. This is of course a well-known pattern, that applies to drugs and vitamins as well. However, there are big differences in exposure conditions, in particular in the control of individual doses, that make the management of hormesis effects of general chemicals very different from the management of therapeutic effects of pharmaceuticals. A demonstrated positive effect at a low level of exposure should of course have impact on exposure limits and other regulations, so that reductions in exposure are not required when such reductions are known to have predominantly negative effects on the exposed individual. At least in many cases, such adjustments seem to be possible through changes in the detailed regulations for specific substances and specific exposures, without changes in the overall structure of current regulatory frameworks.

2. What ethical principles are most appropriate for guiding policy in the face of scientific uncertainty, and how might those principles apply to the hormesis case? For example, would some formulation of the precautionary principle be ethically advisable in this case or not, and what would its ramifications be?

There are two major types of error that you can make in a scientific statement (Hansson 1995). Either you conclude that there is a phenomenon or an effect although it does in fact not exist. This is called an error of type I (false positive). Or you miss an existing phenomenon or effect. This is called an error of type II (false negative).

In the internal dealings of science, errors of type I are the serious errors. To make such an error means to draw an unwarranted conclusion, to believe something that should not be believed. Such errors lead us astray, and if too many of them are committed then scientific progress will be blocked by the pursuit of all sorts of blind alleys.

Errors of type II, on the other hand, are much less serious from a (purely) scientific point of view. To make such an error means that you keep an issue open instead of adopting a correct hypothesis. Of course, not everything can be kept open, and science must progress when there are reasonable grounds for (provisionally) closing an issue. Nevertheless, failing to proceed is in this context a much less serious error than walking in the wrong direction.

This difference in severity between the two types of error can also be expressed in terms of burdens of proof. When determining whether or not a scientific hypothesis should be accepted for the time being, the onus of proof falls to its adherents. Those who claim the existence of an as yet unproven effect – such as a (positive or negative) health effect of a chemical substance – have the burden of proof.

All this applies, as I said, to the internal dealings of science. The situation is different when science is applied to practical problems, for instance in risk assessment. The major difference is that type II errors tend to be much more serious in practical applications.

Often the same or very similar questions are asked in a (purely) scientific context and in a risk assessment context. We can for instance ask the question "Is the fruit of the bog bilberry poisonous?" as a purely scientific question. Then the intra-scientific burden of proof applies in the way that I just described. If the same question is followed by "My four-year old has picked a lot of them and wants to eat them now", then the burden of proof is, presumably, distributed differently.

This example illustrates a general pattern. It would not seem rational – let alone morally defensible – for a decision-maker to ignore all preliminary indications of a possible danger that do not amount to full scientific proof. We typically wish to protect ourselves against suspected health hazards even if the evidence is much weaker than what is required for scientific proof. Therefore, in order to guide the type of decisions that we want to make, risk assessments have to be

based on criteria for the burden of proof that differ in many cases from those used for intrascientific purposes.

It is important to note that this does not mean that we operate with different criteria of truth in the two contexts. Instead, the difference concerns our criteria for reasonable recommendations for action (Hansson 1997, 1999). When acting in intrascientific contexts, the scientist in our example should not act *as if* it is known to be true that these fruits are poisonous. She should not, for instance, write in a textbook that they are poisonous, or refrain from investigating whether they are toxic with the motivation that they are already known to be so. In contrast, the parent is well adviced to act *as if* it is known to be true that these fruits are poisonous: She should make sure that the child does not eat them.

This is not a new insight. It was, for instance, expressed with excellent clarity by the American philosopher Richard Rudner (1953). Today, it is often called the precautionary principle (Sandin 1999).

How does all this relate to hormesis? Here the situation is the opposite of what we are used to in environmental issues: The contested phenomenon is a positive effect, not a negative one. The precautionary principle gives precedence to avoiding the most negative scenario. Therefore, the most direct transferal of the precautionary principle to hormesis would consist in applying stricter requirements of evidence before one acts upon the hypothesis that hormesis exists than before acting on a hypothesis that an adverse effect exists.

However, it does not necessarily follow from this that a hormesis hypothesis should be subjected to stricter standards of proof than those that are used for internal purposes of science. The internal scientific criteria put the burden of proof on the standpoint that an hormesis effect exists rather than on the standpoint that it does not exist. This is also essentially what the precautionary principle requires. For positive effects such as hormesis effects, contrary to negative effects, the requirements of the precautionary principle seem to roughly coincide with ordinary scientific standards of evidence.

3. What ethical considerations should guide public policies that balance potentially positive health impacts of toxins on some individuals against potentially negative health impacts of the same toxins on other individuals? What about balancing positive effects on some biological endpoints against negative effects on other endpoints?

The weighing of risks against benefits is a central feature in a large number of social practices. However, it is performed in fundamentally different ways in different application areas. The crucial difference concerns whether or not benefits for one person are allowed to outweigh harms to another person. We can distinguish between two principles for the weighing of risks against benefits (Hansson 2004, 2006a).

The collectivist risk-weighing principle:
An option is acceptable to the extent that the sum of all individual risks that it gives rise to is outweighed by the sum of all individual benefits that it gives rise to.

The individualist risk-weighing principle:
An option is acceptable to the extent that the risk to which each individual is exposed is outweighed by benefits for that same individual.

Hence, in individualist risk-weighing risks and benefits pertaining to one and the same person can be weighed against each other, whereas risks and benefits for different persons are treated as incomparable.

In risk analysis, the collectivist principle dominates. In the standard calculations of risk analysis as applied for instance to energy production and various other industrial applications, a disadvantage to one person can always be compensated by an equally sized advantage to another person. Just as in classical utilitarianism, individuals have no other role than as carriers of utilities and disutilities, the values of which are independent of whom they are carried by.

The individualist risk-weighing principle dominates in areas emanating from clinical medicine. Dietary advice is one example. Due to environmental contamination, health authorities recommend limits in the consumption of fish caught in certain waters. Such recommendations are based on endeavours to balance the negative health effects of the contaminants against the positive effects of fish as a constituent of the diet (Knuth et al 2003). This balance is struck separately for each individual; thus positive effects for others (such as the fishing industry) are excluded from consideration. Similarly, in standard medical research ethics, each person's participation in a clinical trial has to be defensible according to reasonable assessment of risks and benefits for that particular person. Advantages to future patients cannot outweigh serious risks to the participating patient (Hansson 2006b).

Generally speaking, individualist risk-weighing dominates in those areas of preventive health that have grown out of medical practices, whereas collectivist risk-weighing dominates in many other areas. The health effects of chemical substances in the atmosphere are in general evaluated in terms of collectivist risk-weighing, whereas those of substances in food are evaluated in terms of individualist risk-weighing. This difference seems to have its origins in different social and scientific traditions. It is not obvious how differences such as this can be reconstructed with reference to consistent underlying principles of preventive health or social priority-setting.

Which of these risk-weighing principles should be applied to hormesis effects of environmental pollution? The choice is far from obvious. On one hand, since the effects of environmental pollution are usually discussed in terms of collectivist risk-weighing, it would seem natural to apply these criteria to hormesis as well. On the other hand, hormesis effects are positive health effects, and other positive health effects, such as those of pharmaceuticals, exercise and certain foodstuffs, are usually discussed in terms of individualist risk-weighing.

My personal view is that collectivist risk-weighing of hormesis effects has small chances of gaining general social acceptance. Large segments

of the public would not accept an exposure that is supposed to have positive health effects on others but negative effects on themselves. If that is right, then it is advisable to tailor risk and benefit assessments of hormesis effects to individualist risk-weighing. This means that risk assessors should investigate effects on fairly fine-grained subgroups of the population, rather than contenting themselves with estimates of total population effects.

4. What sorts of scientific investigations would be most helpful for facilitating ethically informed decisions about applying hormesis to public policy?

I would like to point out three types of scientific investigations that are relevant in this context.

1. Most obviously, we need scientific evidence that the effects exist, and documentation showing at what exposure levels they dominate over toxic effects.

2. We need scientific studies of additive and synergy effects for the adverse effects of substances with similar or related modes of action. It is for instance quite possible that certain non-genotoxic carcinogens add to each other's effects. Then, even if there is a threshold for each such substance, an individual whose exposure is below the threshold for each of these substances may nevertheless be at risk due to the combined effects of many substances in her (natural and man-made) environment. If this is so, then an additional exposure may lack a practical threshold even if it has a biological one. As an example, a substance that induces cell proliferation does not have a *practical* threshold if its effects are added to those of other agents that already cause cell proliferation. In that case, a hormesis effect of the same substance has to be weighed against a negative effect that is above the threshold.

3. Exposures. Exposure analysis is often the weakest link in a risk analysis. This can be so in "benefit analysis" as well. Exposure analysis of environmental pollutants is mostly focused on identifying the highest exposures. It is quite another matter to identify the conditions under which there are dose levels at which hormesis effects dominates over adverse effects. This may require the development of new methodology for exposure analysis.

5. How important is it to integrate deliberative modes of public participation into discussions about hormesis and regulatory policy? If deliberation is warranted, what are some promising mechanisms and guidelines for implementing it? What are the responsibilities of scientific researchers when they provide information to deliberative bodies, policy makers, and the public under conditions of scientific uncertainty?

Modern risk analysis is largely based on a quantitative methodology that is, from a decision-theoretical point of view, essentially an application of expected utility maximization (or expected disutility minimization). The severity of a risk is measured as the probability-weighted severity of the negative outcomes that the risk refers to. Hence, a risk characterized by a probability p of a negative event with severity s has the same impact in the calculations as a negative event whose severity equals ps and about which we are certain that it will occur. Beginning with the influential *Reactor Safety Study* (WASH-1400, the Rasmussen report) from 1975, risk has often not only been *measured by* but also *identified with* expected disutility (Rechard 1999). In other words, risk is defined as the product of probability and severity.

Probabilistic risk analysis is a highly useful tool that provides risk managers with important information. However, it does not provide them with all the information that they need in order to make risk management decisions. In particular, important ethical aspects are not covered in these forms of risk analysis. Risks do not just 'exist' as free-floating entities; they are taken, run, or imposed. Risk-taking and risk imposition involve problems of agency and interpersonal relationships that cannot be adequately expressed in a framework that operates exclusively with the probabilities and severities of outcomes (Hansson 2003). In order to appraise an action of risk-taking or risk imposition from a moral point of view, we also need to know who performs the action and with what intentions. For instance, it makes a moral difference if someone risks her own life or that of somebody else in order to earn a fortune for herself. It also makes a difference if risk-taking is freely chosen by the affected person or imposed against her will. Therefore, traditional quantitative analysis of risk needs to be supplemented with a systematic characterization of the ethical aspects of risk, including issues such as voluntariness, consent, intent, and justice.

As the experience with fluoridation of drinking water shows, it is difficult enough to gain public acceptance for exposure that is calculated to have positive health effects on each exposed individual. In order to be useful for public deliberations, an ethical appraisal of hormesis effects will have to deal not only with aggregated effects on the total population but also with the risk-benefit balance on the individual level.

REFERENCES

Hansson, Sven Ove (1995) "The Detection Level", Regulatory Toxicology and Pharmacology 22:103-109.

Hansson, Sven Ove (1997) "Can we reverse the burden of proof?", Toxicology Letters 90:223-228.

Hansson, Sven Ove (1999) "Adjusting Scientific Practices to the Precautionary Principle" Human and Ecological Risk Assessment, 5:909-921.

Hansson, Sven Ove (2003) "Ethical criteria of risk acceptance", Erkenntnis, 59:291-309.

Hansson, Sven Ove (2004) “Weighing Risks and Benefits”, Topoi, 23:145-152.

Hansson, Sven Ove (2006a) “Economic (ir)rationality in risk analysis”, Economics and Philosophy, 22:231-241.

Hansson, Sven Ove (2006b) “Uncertainty and the Ethics of Clinical Trials”, Theoretical Medicine and Bioethics, 27: 149–167.

Knuth, Barbara A., Nancy A. Connelly, Judy Sheeshka, and Jacqueline Patterson (2003) “Weighing Health Benefit and Health Risk Information when Consuming Sport-Caught Fish”, *Risk Analysis* 23:1185-1197.

Rechard, R. P. (1999) Historical Relationship Between Performance Assessment for Radioactive Waste Disposal and Other Types of Risk Assessment. *Risk Analysis* 19: 763-807.

Rudner, Richard (1953) “The scientist qua scientist makes value judgments”, *Philosophy of Science* 20:1-6.

Sandin, Per (1999) “Dimensions of the Precautionary Principle”, *Human and Ecological Risk Assessment* 5: 889-907.

THE HORMESIS CONCEPT AND RISK ASSESSMENT: ARE THERE UNIQUE ETHICAL AND POLICY CONSIDERATIONS?

George R. Hoffmann
Department of Biology
College of the Holy Cross
Worcester, MA 01610-2395
Telephone: 508-793-3416; Fax: 508-793-2696
E-mail: ghoffmann@holycross.edu

William E. Stempsey
Department of Philosophy
College of the Holy Cross
Worcester, MA 01610-2395

ABSTRACT

The hormesis concept holds that low doses of toxic substances and radiation elicit modest biological responses opposite to those caused by higher doses of the same agents. This concept stands in contrast to the prevailing views that a threshold model predicts most responses to toxicants at low doses and that linear extrapolation best predicts mutagenic and carcinogenic responses. Beyond the scientific considerations, there has been concern that inclusion of the hormesis model in risk assessment would raise complex ethical questions, pose serious challenges for policy-makers, and threaten public safety. This article briefly reviews the growing evidence for hormesis and offers a perspective on the related ethical and societal issues. Complexities stem from the nature of biphasic curves, in which biological responses fall both above and below background levels. The monotonic responses of the threshold and linear models lend themselves to a single policy objective – avoiding harm associated with exposures. The biphasic responses of the hormesis model, however, suggest the possibility of accruing benefit as well as avoiding harm. The prospect of applying the hormesis model to public-health policy is impeded by insufficient ability to identify the hormetic and toxic zones with precision. Moreover, heterogeneity among individuals in susceptibility to toxicants suggests that benefit and risk may be distributed unequally in the population. The potential shift in policy objectives associated with hormesis is considered relative to the difficulty of balancing the ethical principles of non-maleficence and beneficence while maintaining a higher priority on the former.

CONTRASTING DOSE-RESPONSE MODELS

Risk assessment for toxic substances and radiation is typically concerned with estimating human health risks at low exposure levels on the basis of laboratory studies at higher doses and epidemiologic data. Understanding the shape of the dose-response curve in the low dose region is therefore critical. Unfortunately, determining the nature of responses at low doses is not straightforward, owing to a lack of statistical power in measuring small differences and quantifying effects occurring at low frequencies. Many data sets are compatible with more than one low-dose model because of random variation in biological responses[1].

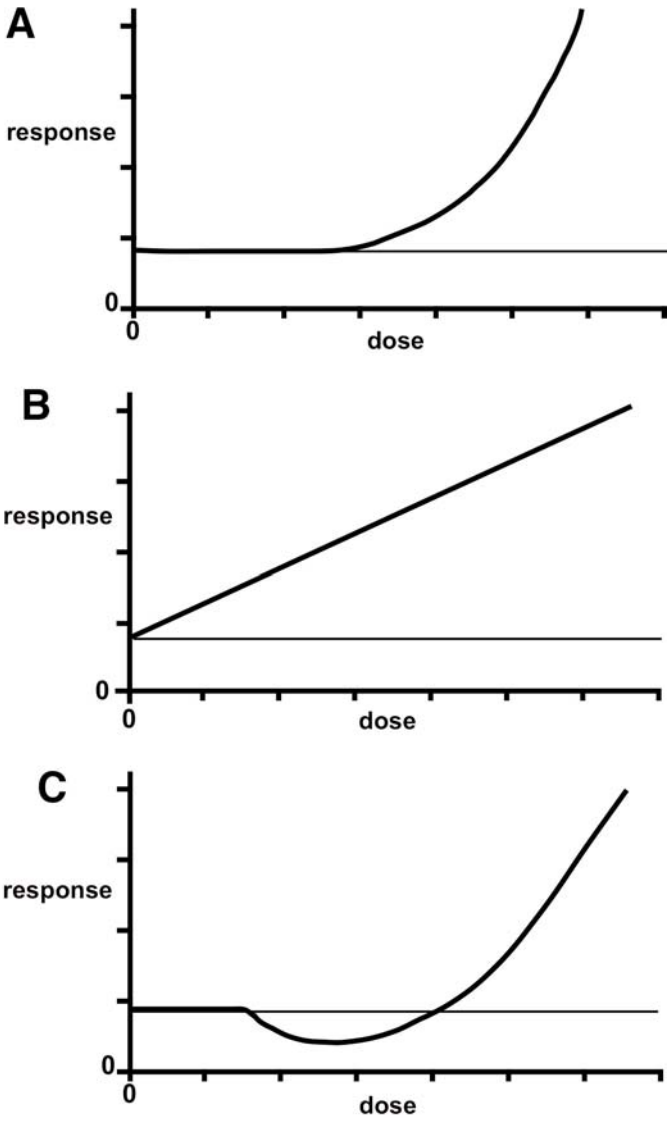

Figure 1. Comparison of dose-response curves described by a threshold model (A), a linear model (B), and the hormesis model (C). Each curve is shown in comparison to the spontaneous frequency of the toxicologic effect being measured.

Figure 1 illustrates dose-response curves described by the threshold, linear, and hormesis models. While the thinking of toxicologists and

policy makers has been dominated by the linear and threshold models, growing evidence suggests that responses to low doses may exhibit hormesis, a phenomenon defined by the effect at low doses being opposite to that elicited by higher doses. The linear and threshold responses are monotonic, in that there is an increase or decrease over the entire range of doses at which there is an effect. In contrast, the hormetic response is biphasic, often described as J-shaped or U-shaped[2,3].

There is nothing inherent in the hormesis concept that says whether effects are beneficial or adverse, only that the effects at low doses are opposite to those occurring at high doses. In the context of toxicology, however, the effects being measured at high doses are typically detrimental. The x axis in Figure 1 is therefore an adverse biological response or dysfunction (e.g., cancer), and the response in the hormetic zone is a frequency of this effect lower than the background frequency (i.e., lower than the frequency occurring spontaneously in the absence of exposure). The same concepts apply to inhibition or loss of a normal biological function. In this case, threshold or linear responses would be monotonic declines (e.g., in growth rate or survival), whereas a hormetic curve would be an inverted U, in which low doses give above-control responses before the decline associated with toxicity.

STATUS OF THE MODELS

The threshold model, which takes the form of a sigmoid curve when applied to proportions of a study population exhibiting a quantal characteristic (e.g., lethality), is widely regarded as the most basic dose-response relationship in toxicology[4]. Unlike most of toxicology, the prevailing model for mutagenesis and carcinogenesis has been linear extrapolation to low doses[5-8], stemming from the view that mutations result from a direct interaction between the agent and its target, following one-hit kinetics[5]. Radiation carcinogenesis at doses for which responses are readily measured is largely compatible with linearity[6]. Nevertheless, for carcinogenesis, as well as other endpoints, the measurement of responses becomes problematic at increasingly low doses. Therefore, acceptance of the view that responses extrapolate monotonically back to zero resides more on theory than on systematic measurement of effects at very low doses. Hormesis has been an unwelcome intruder into a concept of adverse biological effects that is made coherent by linear and threshold models in which all effects are on one side of the control — either above or below, depending on the nature of the endpoint. The hormesis model proposes biological effects both above and below control levels, depending on the dose.

EVIDENCE FOR HORMESIS

An accumulation of evidence argues that hormesis is a real phenomenon, but substantial argument remains with respect to its generality and mechanisms. Hormesis is not obvious in most dose-response curves because it tends to be a modest effect, not easily distinguished from a control, and it occurs at low doses, whereas most studies focus on higher doses that elicit readily observable adverse effects. Effects that occur at low background frequencies (e.g., some tumors) are problematic, as it is impractical to measure a decrease in their frequency[9]. Hormesis is only discernible when a parameter can be measured in two directions — both above and below the control.

Early reports of hormesis, extending to the nineteenth century[2,10], uncovered the general nature of the phenomenon, but they did not reveal its prevalence or adequately explore alternative explanations. Recent studies have relied on extensive, systematic surveys of toxicological literature using carefully delineated criteria. The outcome of such surveys is that biphasic dose responses, fitting the hormesis concept, are common, and some 3000 examples of hormesis had been reported in the scientific literature by 2001[2,9,11]. The hormetic responses tend to be modest, usually differing from the control by less than 50%. They are typically observed immediately below the NOAEL (no observed adverse effect level), and they extend to several-fold, sometimes many-fold, lower doses[2,12].

There has been progress in addressing important criticisms raised by Crump[13] that the existence of many examples of hormesis does not demonstrate its generality if one does not know the denominator that defines its prevalence relative to other response patterns and that information is needed on the frequency of false positives in the identification of responses as hormetic. In a survey by Calabrese and Baldwin[11] of over 20,000 articles, 19.5% of sub-NOAEL doses differed from the control in the hormetic direction, 80% did not differ from the control, and 0.6% differed in the same direction as toxic responses. Thus, responses suggesting hormesis were 32-fold more common than responses not expected for hormesis[11]. Hormetic responses were observed in diverse organisms and encompassed many toxic agents and biological endpoints. Subsequent analyses confirmed and extended these observations[14]. Studies of chemical effects in 136 tumor cell lines also revealed many instances of hormesis and gave insight into general attributes of hormetic curves[12]. Hormesis outperformed the threshold model in an analysis of data from a National Cancer Institute (NCI) drug-screening database containing 56,914 dose responses for 2189 chemicals in 13 yeast strains[15]. In this database, responses consistent with hormesis occurred four times more frequently than expected by chance[15]. Moreover, there is a growing appreciation for mechanisms that can explain hormetic effects[3].

In pointing out the growing evidence for hormesis, we would be remiss if we did not acknowledge that certain phenomena have been reported that, although incompletely understood and of uncertain relationship to hormesis, may influence effects at low dosages. Among these, bystander effects and genomic instability are of special interest. There is evidence that irradiated cells signal nearby nonirradiated cells ("bystanders") and that the latter may experience cytotoxicity, apoptosis, chromosome aberrations, and other effects as a consequence[16-24]. In the case of genomic instability, the effects may be perpetuated in the progeny of the cells initially affected[18,19,22-24]. Bystander effects are most important in the low-dose range, where they are not overshadowed by direct effects of radiation[24]. Although most studied for irradiation, bystander phenomena may extend to chemical exposures, but this is technically difficult to measure[22,23]. It

has been suggested that bystander effects and induced genomic instability, in which nontargeted cells share in adverse effects, may cause deviations from low-dose linearity different from those predicted for hormesis, that is, toward higher risk[6,17]. This interpretation is speculative, and it depends on the doses that induce bystander effects and genomic instability, whether there are thresholds in their induction, how the effects are propagated, and whether beneficial or detrimental effects are more prevalent in the bystander population[21-23]. Such findings hint that hormesis may not be the last surprise in biological responses to low doses of toxicants and radiation. The recognition of phenomena whose implications for risk are still unclear argues for caution in attempting to predict low-dose effects from responses at high doses, even in light of the growing evidence for hormesis.

THE HORMESIS CONTROVERSY

The controversy about hormesis exists both at the scientific and ethical level. The former considers whether the phenomenon of hormesis is real and, if so, whether it is an oddity observed in isolated instances or is typical of biological responses, as might be expected of an evolutionarily conserved response to stress. The latter considers whether hormesis should influence policy judgments related to health risks associated with low-dose exposures. The intensity of the controversy is driven by the ethical level and the societal implications of the claim that low-dose exposures to toxicants or radiation may be beneficial. This claim runs counter to the assumptions underlying regulatory policy under diverse statutes.

If biological responses commonly display biphasic patterns, we may be using incorrect models when we let threshold responses, and especially linear responses, shape our view of what is likely to happen at low doses. Policies based on an incorrect model have a shaky foundation, yet the policies themselves may be in the public interest, in that errors tend to be made on the side of safety — overestimating rather than underestimating risks. An outgrowth of this tension is a tendency of some critics of hormesis[25,26] to deny the reality of hormesis or to conflate science and policy. While the sociology of science is too complex to neatly separate science from the social factors and value judgments that go into scientific practice, we would argue that risk assessment practices should be based on the best possible scientific analysis.

MATURATION OF THE HORMESIS CONCEPT

For the purposes of this discussion, we define hormesis as a dose-response relationship in which low doses elicit a biological response opposite to that caused by high doses. Hormesis, however, is a multifaceted phenomenon with diverse manifestations, and it overlaps with phenomena typically defined on grounds other than shapes of dose-response curves, including adaptive responses[27,28] and preconditioning[29]. The varied treatments of hormesis and related phenomena in the scientific literature have undoubtedly generated confusion as to what hormesis encompasses[30]. Despite its long history, the hormesis concept has emerged in a modern context only recently, and it does not yet have the clarity that comes with scientific maturation.

Elliott[31] has argued that epistemological debates about the existence of hormesis are hindered by conceptual confusion about hormesis. He finds that "hormesis," as used in the current literature, actually includes seven different concepts. Three of these he calls "operational," which means that they are defined in terms of "criteria of application," that is, some biological endpoint. Another three are "mechanistic," involving the isolation of a system in which hormesis is produced through interaction of parts of the system according to causal laws. The seventh concept is "adaptive," holding that hormesis has developed as an adaptive response to biological stressors. Elliott[31] argues that none of these conceptions is adequate for providing epistemological justification for hormesis, but that the mechanistic concepts provide the strongest basis for research that can substantiate or refute the existence and generality of hormesis. This may be the case in terms of gaining a scientific understanding of hormesis. Nonetheless, for public policy considerations the operational conceptions, i.e., predicting outcomes for defined endpoints at low doses, may be more important. Medical practice includes the use of many drugs that are regulated for effectiveness and toxicity without a complete understanding of their mechanisms of action. While understanding mechanisms of action may provide a surer epistemological grasp, established biological endpoints might be sufficient for policy-making purposes.

CHALLENGES FOR RISK ASSESSMENT

The concept of hormesis has been slow to achieve acceptance for a combination of biological and historical reasons[2,10]. However, the crux of the controversy over hormesis lies in the fact that hormesis has implications for risk assessment that many find troublesome. In its simplest form, the hormesis concept suggests that low doses of a toxicant or radiation would pose less risk than predicted using a linear model and might have beneficial effects not predicted by a threshold model. Some find this interpretation dangerous, in that it could lead to a weakening of standards for environmental protection and the protection of public health against small but real risks at low doses.

According to the threshold model (Figure 1A), there is a dose, the threshold, below which there is no effect, either adverse or beneficial. Under experimental conditions, the threshold is approximated by a NOAEL, the estimation of which is limited by the statistical power of the experiments. For purposes of risk assessment, the threshold model lends itself to one objective — avoiding harm. The challenge is to ensure that exposures are below the threshold. In order to do so, a safety factor may be imposed, building in a margin of safety below the estimated threshold. The controversy over hormesis is probably most intense for cancer risk assessment, where the assumption of a linear model (Figure 1B) with no threshold has dominated high-to-low-dose extrapolations[6,8]. Linear models have been favored for mutagenesis and carcinogenesis, not only for the scientific reason that they were compatible with the early data and traditional hit theory[5,7],

but also because they would lead to errors being made on the side of safety. The latter is a policy reason rather than a scientific reason, and the National Academy of Sciences' Committee on the Biological Effects of Ionizing Radiation acknowledged that linear extrapolation was wise for "reasons of prudent conservatism," even though the "statistical error is too large for this expectation to be tested with any rigor"[7]. Decisions based on a linear model, like those based on a threshold model, still hold a single objective — the avoidance of harm. In this case, however, the aim is minimizing harm, in that every increment of dose above background is seen as carrying a finite risk. Since there is no absolutely safe dose, one must extrapolate downwardly to a dose whose risk can be regarded as negligible or acceptable. At first glance, it may seem that favoring an overestimation of risks is satisfactory or even advantageous, in that the error may lead to economic costs for minimizing exposures but not threaten public health. This may be an oversimplification, however, if an exaggeration of perceived risks hinders the development or use of valuable diagnostic or therapeutic procedures, as has been suggested in the Joint Report of the French Academies in its argument against linear nonthreshold extrapolation for ionizing radiation[32].

The hormesis model (Figure 1C), unlike the linear and threshold models, offers the possibility of two objectives — avoidance of harm associated with the toxic zone, and reaping the benefit of the hormetic zone. If the hormesis concept were to be used as a basis for policy in a way that attempted to derive benefit from the hormetic zone, it would necessitate a corresponding shift in the ethical principle behind public policy. Current policy clearly has non-maleficence as its basis. This ethical principle finds its roots in the medical aphorism *primum non nocere* (above all, do no harm), an imprecise translation from the first book of the Epidemics of the Hippocratic corpus. The Hippocratic author of this work actually urged the physician both to do good and avoid harm. Many types of ethical theory today make the same distinction between beneficence, actively doing good, and non-maleficence, avoiding doing harm, but give ethical priority to the latter.

In his well-known theory of duty, for example, Ross[33] distinguishes beneficence and non-maleficence and argues that non-maleficence is "a duty of more stringent character," and "*prima facie* more binding." Even theorists who include both non-maleficence and beneficence in a single ethical principle make a distinction between these two ideas. Frankena[34] identifies four elements in his single principle: 1) One ought not to inflict evil or harm; 2) One ought to prevent evil or harm; 3) One ought to remove evil or harm; 4) One ought to do or promote good. These elements are to be taken serially; that is, the earlier ones take precedence over the later ones. Hence, even in a one-principle theory, non-maleficence is seen as more binding than beneficence. Beauchamp and Childress[35], in their well-known work on principles in biomedical ethics, argue that it is not always the case that non-maleficence is more binding; for instance, the obligation to rescue a subject injured in research may be more binding than the obligation not to harm the subject in the first place. Nonetheless, they admit that in cases where the two principles conflict, non-maleficence is "typically overriding."

If the hormetic range of benefit for particular substances could be precisely delineated, public policy advocating exposure in this range would signal a shift from the non-maleficence that characterizes current public health policy with respect to toxicants, to the beneficence that characterizes other more controversial public health policies such as vaccination and fluoridation of water supplies. Some advocates of hormesis have recognized this shift. Cook and Calabrese[36], for example, note the difference between "protect," the "attempt to maintain the frequency of disease near background," and "promote," which aims at "reducing the frequency of disease below background, i.e., improving the health of the general public." They argue that the hormetic model will allow decision makers not only to "protect" health but to "optimize" health. What is not explicitly stated, however, is that this reflects a signal change in the ethical principles underlying public policy regarding toxicants: the addition of the principle of beneficence to the traditional emphasis on non-maleficence. It remains to be seen whether the concept of hormesis can be clarified and empirical studies refined enough to define the precise range of benefit for different substances without introducing a risk of harm.

QUANDARIES OF HORMESIS

Rejection of the hormesis model, if the model is correct, implies that beneficial effects of low-dose exposures might be lost. One might argue that this, in itself, is a form of harm, but incurring risks in trying to harvest a benefit runs counter to the principle of non-maleficence and the widely accepted view that regulatory practices should make their errors on the side of safety. Even if one accepts the growing evidence that hormetic responses are more prevalent than those predicted by the threshold and linear models, one need not conclude that regulating to the hormetic zone should be the basis for policy. Hormesis may be sufficiently prevalent to be a default assumption for the scientific interpretation of effects at low doses[15], but it may be premature to use it as a basis for setting exposure standards in risk assessment. While a coherent argument can be made for allowing a precautionary principle to influence regulatory decisions, doing so should be a deliberate policy decision, not an adherence to an incorrect dose-response model as though it were true. Denying the growing evidence for hormesis does not protect public health.

Attempting to regulate to the hormetic zone would necessitate greater precision in identifying the NOAEL than is required for risk assessment using a threshold model with a substantial safety factor. Quantitatively, hormesis is a modest effect, and the benefit is consequently small relative to the risk of accidentally being in the toxic zone where adverse responses can be substantial. If hormesis is indeed general, it may still be reasonable to conclude that it is too dangerous to attempt to harvest its beneficial effect. Potential pitfalls lie in the difficulty of gauging effects at low doses precisely, heterogeneity in susceptibility to toxicants, differences among biological endpoints with respect to the minimal exposures causing adverse effects, specific responses that do not exhibit hormesis, interactions among agents, the potentially high cost of errors, and many other factors that affect responses to toxicants. Even accepting the evidence of hormesis, one may find the scientific uncertainty affecting our ability to identify the

hormetic zone too great to be confident of avoiding the toxic zone for some effects or some individuals. This view argues against regulating to the hormetic zone today, but it leaves open the possibility of doing so in the future if the supporting evidence were sufficient.

Genetic heterogeneity in susceptibility to toxicologic effects has long been recognized, and public health policies need to accommodate it. Those concerned about the implications of hormesis have argued that the failure to account for such heterogeneity is a great failing of the hormesis concept[25]. In fact, genetic heterogeneity does not argue against the hormesis concept, but it would need to be carefully considered if one were to use the hormesis concept as a basis for regulation, just as it does assuming any other model. An analysis of published literature[37] suggests that individuals at high risk and sensitive species typically (but perhaps not always) display hormetic responses, but the curves are shifted to the left on the dose-response scale.

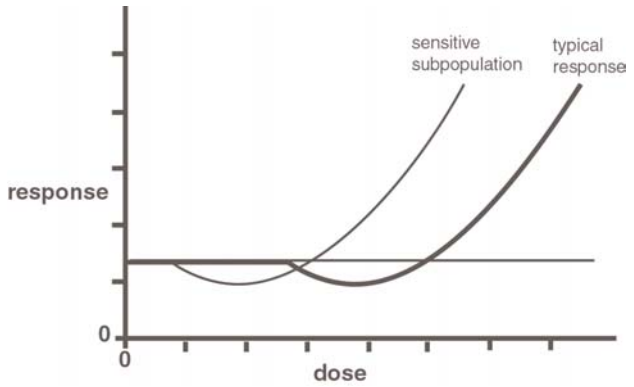

Figure 2. Hypothetical curves showing hormetic responses of a typical population and a genetically sensitive subgroup. The response represents the frequency of an adverse effect within each subpopulation. The thickness of the curves symbolizes the fact that the sensitive subpopulation is a small fraction of the total population.

Figure 2 shows two hypothetical hormetic curves, one for the general population and another for a sensitive subgroup that represents a small fraction of the total population. If the hormetic curves do not overlap, then doses that are hormetic for the population as a whole are toxic for the sensitive subgroup. Regulating in this circumstance raises the specter of part of the population being harmed, while another part benefits. As such, it poses ethical questions not inherent to monotonic responses. Even if NOAEL's could be measured precisely in both populations, there remains uncertainty about how to achieve an optimal solution for both subpopulations. The spectrum of possible responses may be defined by two polar extremes. At one pole is the guiding principle of "do no harm." At the other is a course of utilitarian ethics, in which one seeks the greatest good for the greatest number. In the first instance, benefit would be lost to the majority in order to protect the minority from harm. In the second, a lower incidence of total adverse effect would be bought at the cost of detriment to the sensitive subpopulation. The primacy of non-maleficence argues for steering closer to the former.

In actuality, one could envision a continuum between these poles. The experience of clinical trials may offer a parallel. If serious adverse effects turn up early in the course of a trial, there is a strong tendency to abort the trial, despite the fact that doing so may entail foregoing significant benefit to the majority of the population. This tendency is offset to an extent by a reluctance to terminate the trial prematurely on insufficient grounds. If one attempted to regulate to the hormetic zone, the challenge might be in finding a position sufficiently to the left on the exposure scale to ensure protection of sensitive subgroups in the population. It could be argued that Figure 2 is the worst-case scenario for regulating on the basis of hormesis, in that the hormetic zones of the sensitive and typical populations barely overlap. If the curves were closer together on the x axis, a position below the NOAEL for the sensitive population may still be hormetic in the majority population. Cook and Calabrese[36] have suggested on this basis that exposures below the NOAEL for the sensitive subpopulation are apt to be beneficial for both subpopulations. The problems posed by sensitive subgroups in the population are complex, but they may not be intractable if risk assessment were to reach a refined stage characterized by an accuracy and precision much greater than is possible today.

The problem of subpopulations with differing susceptibilities has been discussed by Gaylor[38], who rightly concludes that it would take "considerable data and/or assumptions" to select proper reference doses for the total population. It would seem prudent to endorse his conclusion that no general recommendations can be given for altering doses on the basis of hormesis. Instead, each situation should be considered on a case-by-case basis. Indeed, proponents of hormesis[36] recommend that any public policy decisions to adjust exposure limits would have to be made in a way that is "transparent."

CONCLUSIONS

Growing evidence supports the reality of hormesis. The extent of its generality remains to be established, but it is clearly prevalent in the responses of diverse biological systems to stress, including that imposed by radiation and a broad array of toxicants. In our view, the hormesis concept needs to be evaluated empirically, independently of how it may figure into public-health policy decisions. Even if true and general, hormesis may be rejected as a basis for policy if it were in the public interest to do so, but this does not alter the reality of the phenomenon. The question as to whether hormesis should figure into toxicology policy may need to be revisited repeatedly as the ability to assess risks accurately and precisely improves. Genetic heterogeneity and the need to protect sensitive subgroups have been used here to explore questions posed by biphasic curves. Exposures that are hormetic for one biological endpoint but adverse for another can similarly pose difficult challenges. Regulation of toxicants to the hormetic zone is fraught with difficulties that will need to be evaluated over time. Other problems that need to be considered include ecological effects of toxicants at low doses and effects on sensitive species. We suspect that exploiting the hormetic zone for purposes of medical or agricultural applications may precede the assimilation of hormesis into toxicology risk assessment.

The slow course of paradigm change in toxicology may allow maturation of the hormesis concept[39] so that if it is ultimately incorporated into policy decisions, it is done in a way that fosters public health and safety, environmental protection, and sound ethical principles.

ACKNOWLEDGMENTS

We thank Edward Calabrese and Kevin Elliott for inviting us to contribute to this collection of papers. We have benefited greatly from the insights both of hormesis proponents and of skeptics; we especially thank Edward Calabrese, Barbara Callahan, Ralph Cook, Walter Kozumbo, Donald MacPhee, Kenneth Prestwich, and Michael Shelby for valuable discussions.

REFERENCES

1. Calabrese EJ. Toxicological diversity: making room for the U-shaped dose-response. *Hum Exper Toxicol* 2003; 22: 465-466.

2. Calabrese EJ. Hormesis: Changing view of the dose response; a personal account of the history and current status. *Mutat Res* 2002; 511: 181-189.

3. Conolly RB, Lutz WK. Nonmonotonic dose-response relationships: mechanistic basis, kinetic modeling, and implications for risk assessment. *Toxicol Sci* 2004; 77: 151-157.

4. Eaton DL, Klaassen CD, Principles of toxicology. In: Klaassen CD (ed.). *Casarett & Doull's Toxicology: The Basic Science of Poisons*. 6th edition. New York: McGraw-Hill, pp. 11-34, 2001.

5. Auerbach C. *Mutation Research: Problems, Results, and Perspectives*. London: Chapman and Hall, 1976.

6. Brenner DJ, Doll R, Goodhead DT, Hall EJ, Land CE, Little JB, Lubin JH, Preston DL, Preston RJ, Puskin JS, Ron E, Sachs RK, Samet JM, Setlow RB, Zaider M. Cancer risks attributable to low doses of ionizing radiation: Assessing what we really know. *Proc Natl Acad Sci USA* 2003; 100: 13761-13766.

7. National Research Council (NRC Committee on the Biological Effects of Ionizing Radiations). *The Effects on Populations of Exposure to Low Levels of Ionizing Radiation* (BEIR III). Washington, DC: National Academy Press, 1980.

8. National Research Council (NRC Committee to Assess Health Risks from Exposure to Low Levels of Ionizing Radiations). *Health Risk from Exposure to Low Levels of Ionizing Radiation* (BEIR VII, Phase 2). Washington, DC: National Academy Press, 2006.

9. Calabrese EJ, Baldwin LA. Hormesis: A generalizable and unifying hypothesis. *Crit Rev Toxicol* 2001; 31(4 & 5): 353-424.

10. Calabrese EJ. Historical blunders: How toxicology got the dose-response relationship half right. *Cell Molec Biol* 2005; 51: 643-654.

11. Calabrese EJ, Baldwin LA. The frequency of U-shaped dose responses in the toxicological literature. *Toxicol Sci* 2001; 62: 330-338.

12. Calabrese EJ. Cancer biology and hormesis: Human tumor cell lines commonly display hormetic (biphasic) dose responses. *Crit Rev Toxicol* 2005; 35: 463-582.

13. Crump K. Evaluating the evidence for Hormesis: A statistical perspective. *Crit Rev Toxicol* 2001; 31: 669-679.

14. Calabrese EJ, Baldwin LA. The hormetic dose-response model is more common than the threshold model in toxicology. *Toxicol Sci* 2003; 71: 246-250.

15. Calabrese EJ, Staudenmayer JW, Stanek EJ, Hoffmann GR. Hormesis outperforms threshold model in NCI anti-tumor drug screening database. *Toxicol Sci* 2006; 94: 368-378.

16. Belyakov OV, Mitchell SA, Parikh D, Randers-Pehrson G, Marino SA, Amundson SA, Geard CR, Brenner DJ. Biological effects in unirradiated human tissue induced by radiation damage up to 1 mm away. *Proc Natl Acad Sci USA* 2005; 102: 14203-14208.

17. Morgan WF. Genomic instability and bystander effects: a paradigm shift in radiation biology? *Mil Med* 2002; 167(Suppl. 1): 44-45.

18. Morgan WF. Is there a common mechanism underlying genomic instability, bystander effects and other nontargeted effects of exposure to ionizing radiation? *Oncogene* 2003; 22: 7094-7099.

19. Morgan WF. Non-targeted and delayed effects of exposure to ionizing radiation: I. Radiation-induced genomic instability and bystander effects *in vitro*. *Radiat Res* 2003; 159: 567-580.

20. Morgan WF. Non-targeted and delayed effects of exposure to ionizing radiation: II. Radiation-induced genomic instability and bystander effects *in vivo*, clastogenic factors and transgenerational effects. *Radiat Res* 2003; 159: 581-596.

21. Morgan WF, Sowa MB. Effects of ionizing radiation in nonirradiated cells. *Proc Natl Acad Sci USA* 2005; 102: 14127-14128.

22. Mothersill C, Seymour CB. Radiation-induced bystander effects and adaptive responses — the Yin and Yang of low dose radiobiology? *Mutat Res* 2004; 568: 121-128.

23. Mothersill C, Seymour CB. Radiation-induced bystander effects — implications for cancer. *Nat Rev Cancer* 2004; 4: 158-164.

24. Mothersill C, Seymour CB. Radiation-induced bystander effects and the DNA paradigm: An "out of field" perspective. *Mutat Res* 2006; 597: 5-10.

25. Axelrod D, Burns K, Davis D, Von Larebeke N. "Hormesis" – an inappropriate extrapolation from the specific to the universal. *Int J Occup Environ Health* 2004; 10: 335-339.

26. Thayer KA, Melnick R, Burns K, Davis D, Huff J. Fundamental flaws of hormesis for public health decisions. *Environ Health Perspect* 2005; 113: 1271-1276.

27. Cairns J. Efficiency of the adaptive response of *Escherichia coli* to alkylating agents. *Nature* 1980; 286: 176-178.

28. Wolff S. The adaptive response in radiobiology: evolving insights and implications. *Environ Health Perspect* 1998; 106: 277-283.

29. Murry CE, Jennings RB, Reimer KA. Preconditioning with ischemia: a delay of lethal cell injury in ischemic myocardium. *Circulation* 1986; 74: 1124-1136.

30. Calabrese EJ, Bachmann KA, Bailer AJ, Bolger PM, Borak J, Cai L, Cedergreen N, Cherian MG, Chiueh CC, Clarkson TW, Cook RR, Diamond DM, Doolittle DJ, Dorato MA, Duke SO, Feinendegen L, Gardner DE, Hart RW, Hastings KL, Hayes AW, Hoffmann GR, Ives JA, Jaworowski Z, Johnson TE, Jonas WB, Kaminski NE, Keller JG, Klaunig JE, Knudsen TB, Kozumbo WJ, Lettieri T, Liu SZ, Maisseu A, Maynard KI, Masoro EJ, McClellan RO, Mehendale HM, Mothersill C, Newlin DB, Nigg HN, Oehme FW, Phalen RF, Philbert MA, Rattan SI, Riviere JE, Rodricks J, Sapolsky RM, Scott BR, Seymour C, Sinclair DA, Smith-Sonneborn J, Snow ET, Spear L, Stevenson DE, Thomas Y, Tubiana M, Williams GM, Mattson MP. Biological stress response terminology: Integrating the concepts of adaptive response and preconditioning stress within a hormetic dose-response framework. *Toxicol. Appl. Pharmacol.* 2007; 222:122-128.

31. Elliott K. Conceptual clarification and policy-related science: The case of chemical hormesis. *Perspect Sci* 2000; 8: 346-366.

32. Tubiana M, Aurengo A. Dose-effect relationship and estimation of the carcinogenic effects of low doses of ionising radiation: the Joint Report of the Académie des Sciences (Paris) and of the Academie Nationale de Médecine. *Int J Low Radiation* 2006; 2: 1-19.

33. Ross WD. *The Right and the Good*. Indianapolis: Hackett Publishing Company, pp. 21-22, 1988.

34. Frankena W. *Ethics*. 2d edition. Englewood Cliffs, NJ: Prentice-Hall, p. 47, 1973.

35. Beauchamp TL, Childress, JF. *Principles of Biomedical Ethics*. 5th edition. New York: Oxford University Press, pp. 114-116, 2001.

36. Cook R, Calabrese EJ. The importance of hormesis to public health. *Environ Health Perspect* 2006; 114:1631-1635.

37. Calabrese EJ, Baldwin LA. Hormesis and high-risk groups. *Reg Toxicol Pharmacol* 2002; 35: 414-428.

38. Gaylor D. Safety assessment with hormetic effects. *Hum Exp Toxicol* 1998; 17: 251-253.

39. Elliott K. Error as a means to discovery. *Phil Sci* 2004; 71: 174-197.

RISKS TO HEALTH AND RISKS TO SCIENCE: THE NEED FOR A RESPONSIBLE "BIOEVIDENTIAL" SCRUTINY

Deborah G. Mayo
Department of Philosophy
Virginia Tech
Blacksburg, VA 24061, USA
Phone 540-231-8488
E-mail: mayod@vt.edu

Aris Spanos
Department of Economics
Virginia Tech
Blacksburg, VA 24061, USA
E-mail: aris@vt.edu

INTRODUCTION AND AIMS

We are pleased to take part in this forum on ethical issues of hormesis risk assessment and policy. In our view, ethical issues surrounding evidence-based risk policy in general are not properly addressed if divorced from issues of the responsible interpretation of the associated risk evidence. The former, bioethical issues, are adequately addressed only along with an accompanying methodological critique that may be dubbed "bioevidential". Just as bioethics requires developing and applying knowledge of ethical theory and principles to the assessment of controversial risk policies, bioevidentialism calls for applying a critical understanding of theories of data, statistical modeling, and inference to the evaluation and assessment of controversial risk evidence.

We do not present ourselves as medical or toxicological experts. However, our combined areas of expertise—philosophical foundations of science, statistical inference and modeling—enables the critical evaluation of the uncertainties, assumptions, and errors along the manifold steps in arriving at inductive/statistical inferences underlying risk assessments. The focus here is evidence for hormetic hypotheses concerning carcinogenic risks. Our goal is not to pass judgment on the truth or falsity of hormetic theory, but to evaluate the epistemological warrant of the evidence given in support of hormetic hypotheses by some of their main advocates.

It is laudable that leading hormesis proponents are opening the evidential and policy-laden issues to widespread critical appraisal, as represented by this and other forums. We aim not to provide ammunition to those who take issue with the likely policy implications of accepting hormesis, but to constructively suggest how hormesis proponents may strengthen existing efforts at responsible self-criticism, and in so doing demonstrate the ethical soundness of the evidence on which recommended policies are based. We examine both the evidential sources themselves and critical overviews: Crump (2001), Zapponi and Marcello (2006), Thayer, et al (2005) and Kitchin and Drane (2005). Our remarks are also informed by the American Statistical Association's "Ethical Guidelines for Statistical Practice" which lists such rules as "Report the limits of statistical inference of the study and possible sources of error". As we proceed we will offer constructive suggestions for reporting if not ameliorating such errors in inference. We conclude that the consequences of deciding risk management policy with the current knowledge gaps poses risks not only to health but to science; see Mayo (1991).

A MINIMAL STANDARD FOR EVIDENCE

Hormesis refers to a phenomenon where a substance that is deleterious at high doses causes a response in the opposite direction at low doses (we can call such low dose reversals "improvements" to avoid calling them benefits.) Although some hormetic effects are uncontroversial, existing use of the linear threshold model in toxicology already allows taking these into account (via U or J shaped models) on a case by case basis. Calabrese and Baldwin (C&B) well-known supporters of hormetic theory want to go much further: they claim to have provided sufficient evidence to change the default assumption in toxicology in general. We assume the main claims of Calabrese and Baldwin (e.g., 1998, 2003), Calabrese (2005) are well known to readers of this forum.

Evidence for hormetic hypotheses are based on data that disagree with one or more 'no effect' or null hypotheses asserting:

H_0: there is 0 risk decrease, or 0 improvement, at low doses. (H_0 might also include risk increases.) Although an observed risk decrease in low-dose compared to untreated (controls) does not logically contradict H_0, it may be regarded as statistical grounds for inferring:

H_1: there is evidence of improvements or decreased risk at low doses, which may then be the basis for a *hormesis hypothesis:*

H: observed improvements are due to a hormetic effect.

Data **x** purporting to provide evidence for hormesis, at minimum, accords with H_1 but more is required to have genuine evidence for *H.* Mere accordance with the data is too easy whether for statistical hypothesis H_1 or a substantive hormesis hypothesis *H*.

We focus here on the least stringent standard for evidence: if it can be shown that the observed accordance between **x** and *H* would very probably have occurred even if *H* is false, or if the test turns out to have very poor ability to discriminate between cases where *H* is genuinely indicated by **x** and those where it would be clearly fallacious to infer *H*, then there are grounds to question the scientific credentials of the particular inference to *H*. We can abbreviate this:

> ***Severity principle (Weak):*** If data **x** 'accords with' *H* but the test very probably would have erroneously inferred *H* even if false, then *H* is *not well warranted by* **x**.

To run afoul of this weak severity principle would seem to abrogate the very basis for using empirical data to appraise hypotheses, and is scarcely a source of controversy.

Far murkier are questions about what is required to show that seriously insevere tests are avoided. How does one succeed in inferring only reasonably severely warranted hypotheses? The bioevidentialist program approaches these questions by identifying classic examples of flaws and foibles of general types that are found across the landscape of uncertain inferences, whether formal or informal. If one deliberately considers circumstances that would, with high probability, have told against an observed accordance between data and *H*, and yet no flaw or error is detected, then the severity with which *H* passes is fortified; see Mayo & Spanos (2006a). It is therefore highly advantageous, if not obligatory, for those claiming to have evidence for *H* to show at least that egregious lack of severity is avoided. Bioevidentialist scrutiny can provide systematic ways to check this.

HUNTING FOR SIGNIFICANT HORMETIC EFFECTS IN THE LITERATURE

Calabrese and Baldwin (1998, 2003) obtain their evidence of hormesis through an extensive literature search of existing studies, carried out for different reasons, rather than through controlled trials testing a null hypothesis of no improvement. Since this may well be the only reasonable evidence available at present, it is important to address issues of evidence regarding these literature searches and the uses hormetic proponents make of them.

Among various methodological questions to which these studies give rise, the most notable are questions arising out of the effect of 'hunting for statistical significance'[i]. Although insisting on a low significance level before rejecting H_0 in favor of H_1 ensures a low probability of erroneously inferring evidence of improvement H_1 (low type I error probability), this error probability guarantee breaks down in the case of searching. In the hormetic case, the searching would be for low-doses, or for risk factors, that are prima facie consistent with hormesis. We may refer again to the Ethical Guidelines of the ASA (1999) which stipulates the need to:

> "Recognize that any frequentist statistical test has a random chance of indicating significance when it is not really present. Running multiple tests on the same data set at the same stage of an analysis increases the chance of obtaining at least one invalid result. Selecting the one "significant" result from a multiplicity of parallel tests poses a grave risk of an incorrect conclusion. Failure to disclose the full extent of tests and their results in such a case would be highly misleading."

Let us put the issue as non-technically as possible: In order to avoid insevere inferences to H_1, standard statistical tests direct one to reject H_0 and infer data **x** provide evidence of a risk decrease if and only if the observed risk decrease is statistically significant at a small level α (e.g., .01 or .05). Suppose, however, that one searches through twenty differences and reports just the one that reaches a significance level of .05. The probability of finding at least one, .05 level, nominally statistically significant difference out of 20, even if all the null hypotheses are true, is approximately .64 [i.e., $(1 - .95^{20})$]. So the type I error probability would be .64, not .05. The inference to the non-null alternative H_1 has passed an insevere test. This concern is behind Crump's (2001) remarks:

"In order to properly control for the false-positive rate one would need to know how extensive the search was that located the data set. If the data set was the most hormetic looking out of 100 examined, then to conduct a statistical test for hormesis at the standard 0.05 level one should use p = 0.0005 [the solution to $1–(1-p)^{100} = 0.05$] rather than p = 0.05." (Crump 2001, p. 672).

In other words, one would need to insist on a much smaller significance level for each case examined in order for the overall type I error probability to remain small. The task for the bioevidentialist is not to fix precise significance levels or other error probabilities, but to raise the kinds of problems that can prevent controlling error rates.

The data from the literature search may be all that is reasonably available, but it is important to recognize that they are not a random selection from all relevant studies. C&B have developed a specially designed point system to ferret them out. We discuss some problems with this point system elsewhere (Mayo and Spanos, 2006b). Crump demonstrates a lack of control of the type I error probability by applying their scoring rules to data deliberately generated so that the null hypothesis is true (no hormesis). Such a simulation allows determining what distribution of scores would be expected from studies in which a hormetic effect is not present (i.e., false-positive rate.) Crump finds, based on his simulation, that "Using the same scoring system, between 94.9% and 99.7% of the simulated data sets showed some evidence of hormesis (score > 2), even though no hormetic effect was present." (Crump, 2001, p. 675). However, Crump's charge may be mitigated if this scoring system is merely to pinpoint cases worth following up. Even if many are actually not hormetic, C&B may escape the charge of high type I error rate so long as the cases identified as potentially supplying hormetic evidence are properly treated. We now turn to this.

ARE CRITICISMS MITIGATED?

The relevant criticisms could be mitigated in a number of ways. First, one may insist that the observed improvement picked out for closer scrutiny (by their scoring algorithm) show, in the original study, *a statistically significant* improvement. Second, one can help mitigate selection bias by a deliberate consideration of as much as possible of the available risk evidence, including factors with both increased and decreased risks as well as other studies on the same risks. Third, even failing to mitigate these threats to validity (by the first two means), clearly revealing this, and taking steps to scrupulously avoid misleading claims, would disarm criticisms. However, thus far, the hormetic proponents appear not to have mitigated and rarely fully expose such noteworthy shortcomings.

Improvements are Statistically Insignificant. Questions arise from the fact that the cases with the most impressive hormetic-looking effects have been picked out for close scrutiny precisely because they show a high incidence among controls. By chance alone, from time to time, a control group may show a higher than normal incidence of an effect, and a thorough literature search is bound to find them[ii]. The obvious danger is that the most impressive hormetic looking effects may simply be aberrations. Zapponi and Marcello (2006) point out a number of cases where the apparent evidence for hormesis is explainable by such high controls (despite the pattern reversing in other trials). Moreover, even where the incidence rate among low-treated subjects is lower than controls (else they would not have been picked out), the observed decrease is virtually never statistically significant.

To understand the implications of this, consider what is being asked in probing the relevant null hypothesis: can the hormetic dose group be considered to have come from the same population as the controls (with respect to the incidence of the effect in question)? Evidence for hormesis would correspond to a 'no' answer, and in particular, a no answer that results because the incidence rate in the low-dose group is statistically significantly lower than in the controls. That observed differences are insignificant means they fail to supply evidence against the null hypothesis:

$$H_0: (p_C - p_T) = 0 \text{ versus } H_1: (p_C - p_T) > 0$$

p_C and p_T being the population relative frequencies of the risk effect in the control vs. low-dose treated groups respectively. That the observed differences fail to reach statistical significance says, in effect, that the low dose group may be considered to have come from the same population as the control group. This is evidence against the hormetic effect in question[iii]. This underscores the danger of relying on a point estimate for dose-response without supplying an associated estimate of its reliability (e.g., via a standard error)

Problems also arise as regards generalizability. The many agents or substances that have an incidence rate of zero (0) or close to zero in the control group are omitted from the literature analysis of hormetic effects; see Zapponi and Marcello (2006). C&B (1998) are searching for cases where a low-dose treated group (of rats) show less cases with the risk effect than controls: there would be no room for observed improvement if controls are already 0. Since many substances associated with risk increases have 0 or near 0 risk rates among controls, it may be of concern that positive support for hormesis from the literature search does not extend to them.

Incompleteness of Evidence and Selective Reporting. Unlike deductive inference, where if a set of premises entails a conclusion *H*, then so do these premises in addition to others, in inductive inference, the addition of other premises can easily turn an impressive looking inference into an illicit one. In particular, to assess overall improvement, it must be recognized that substances are often linked with several risk effects. Selectively reporting on improvements, say, a decreased incidence of testicular cancer, when at the same low dose the data show an increased incidence of some other cancer, would be to omit important information; see Thayer et al (2005). Yet the study of the effects of cadmium chloride on the incidence of *testicular* tumors in male rats is taken as a striking example of hormesis while overlooking relevant evidence reported in the same study that cadmium injections at low doses (hormetic effect region) increased significantly the incidence of *prostate* tumors. Waalkes (2003) makes a good case that prostate tumors constitute the more serious effect on health because the testicular tumors are usually benign. When these results are viewed in conjunction with the relevant significance levels, the evidence for beneficial hormetic effects are called into question.

These seem reasonable questions many of which critics have asked. Scientific responsibility would seem to call for direct responses. Acknowledging them up front, will be the best way to disarm critics and strengthen the evidential credentials of the hormetic research program.

WHAT KINDS OF INFORMATION WOULD BE USEFUL?

(1) Reliable estimates of control incidence rates would enable determining if the high incidence among controls that form the most impressive evidence for hormesis are likely to be due to chance, to background exposure, or to unusually high susceptibility in the animals observed.

(2) Rather than ignore cases with 0 incidence in the control, it would be good to check that no increased incidence is seen even at the very low doses being examined. If none is seen, it would fortify the cases purporting to show evidence of hormesis, because it would increase the severity of the analysis. Were it a mere aberration we might expect increased risk incidence with low doses, so to the extent that none are seen, the cases picked out for study are strengthened.

(3) Now that hormetic hypotheses are achieving fairly widespread attention, we think that attempts to carry out genuinely controlled studies, with several gradations in the hormetic range, for at least some of the more impressive looking cases, should be considered. This will enable the researcher to assess the validity of the underlying statistical model in order to ensure the reliability of inductive inferences; see Mayo and Spanos (2004).

REFERENCES

American Statistical Association (1999), "Ethical Guidelines for Statistical Practice", http://www.amstat.org/profession/index.cfm?fuseaction=ethicalstatistics.

Calabrese, E. J. (2005), "Hormetic Dose-Response Relationships in Immunology: Occurrence, Qualitative Features of the Dose-Response, Mechanistic Foundations, and Clinical Implications," *Critical Reviews in Toxicology*, 35: 89-295.

Calabrese, E. J. and L. A. Baldwin (1998), "Can the Concept of Hormesis Be Generalized to Carcinogenesis?" *Regulatory Toxicology and Pharmacology*, 28: 230-241.

Calabrese, E. J. and L. A. Baldwin (2003), "The Hormetic Dose-Response Model Is More Common than the Threshold Model in Toxicology," *Toxicological Sciences*, 71: 246-250.

Crump, K. (2001), "Evaluating the Evidence for Hormesis: A Statistical Perspective," *Critical Reviews in Toxicology*, 31: 669-679.

Elliott, K. C. (2006), "A Novel Account of Scientific Anomaly: Help for the Dispute Over Low-Dose Biochemical," *Philosophy of Science*, 73:790-802.

Kitchin, K. T. and J. W. Drane (2005), "A critique of the use of hormesis in risk assessment," *Human and Experimental Toxicology*, 24: 249-253.

Mayo, D. G. (1991), "Sociological vs. Metascientific Views of Risk Assessment," pp. 249-279 in Mayo and Hollander (1991).

Mayo, D. G. (1996), *Error and the Growth of Experimental Knowledge*, The University of Chicago Press, Chicago.

Mayo, D. G. (2004), "An Error-Statistical Philosophy of Evidence," in M. Taper and S. Lele (eds.), *The Nature of Scientific Evidence: Statistical, Philosophical, and Empirical Consideration,* " Chicago: University of Chicago Press.

Mayo, D.G. and R. D. Hollander (1991), *Acceptable Evidence: Science and Values in Risk Management,* Eds., Oxford University Press, Oxford.

Mayo, D. G. and M. Kruse (2001), "Principles of Inference and their Consequences," pp. 381-403 in *Foundations of Bayesianism*, edited by D. Cornfield and J. Williamson, Kluwer Academic Publishers, Netherlands.

Mayo, D. G. and A. Spanos (2004), "Methodology in Practice: Statistical Misspecification Testing", *Philosophy of Science*, 71: 1007-1025.

Mayo, D. G. and A. Spanos (2006a), "Severe Testing as a Basic Concept in a Neyman-Pearson Philosophy of Induction," *The British Journal for the Philosophy of Science*, 57: 323-357.

Mayo, D. G. and A. Spanos (2006b), "Philosophical Scrutiny of Evidence of Risks: From Bioethics to Bioevidence," *Philosophy of Science*, 73: 803-816.

Mayo, D. G. and D. R. Cox (2006), "Frequentist Statistics as a Theory of Inductive Inference," pp. 77-97 in *Optimality: The Second Erich L. Lehmann Symposium*, vol. 49, Institute of Mathematical Statistics. Lecture Notes-Monograph Series, Ohio.

Thayer, K. A., R. Melnick, K. Burns, D. Davis and J. Huff (2005), "Fundamental Flaws of Hormesis for Public Health Decisions," *Environmental Health Perspectives*, 113: 1271-1276.

Waalkes, M. P. (2003), "Cadmium Carcinogenesis," *Mutation Research*, 533: 107-120.

Zapponi, G. A. and I. Marcello (2006), "Low-Dose Risk, Hormesis, Analogical and Logical Thinking, *Annals of the N.Y. Academy of Sciences*, 1076: 839-857.

FOOTNOTES

[i] For some general discussion see Mayo, 1996, Mayo and Kruse, 2001, Mayo and Cox, 2006.

[ii] Likewise, however, one can find apparent improvements (observed risk decreases) in the highest dosed groups.

[iii] For instance, on the basis of table 1 in C&B (1998), the test statistic comparing the difference between the proportions of the control and treated groups at low dose (.01) in male rats is:

$$\frac{\frac{10}{73}-\frac{6}{71}}{\sqrt{\frac{\frac{10}{73}\left(1-\frac{10}{73}\right)}{73}+\frac{\frac{6}{71}\left(1-\frac{6}{71}\right)}{71}}}=1.008[.157]$$

with a p-value in square brackets. Similar lack of significance can be shown for each entry.

AN ETHICAL APPRAISAL OF HORMESIS:

TOWARDS A RATIONAL DISCOURSE ON THE ACCEPTABILITY OF RISKS AND BENEFITS

Ortwin Renn
University of Stuttgart
Interdisciplinary Research Center for Risk Governance and Sustainable Technology Development (ZIRN)
Seidenstr. 36
70174 Stuttgart, Germany
Phone: +49-711-6858-3970; Fax: +49-711-6858-2487
E-mail: ortwin.renn@sowi.uni-stuttgart.de

1. INTRODUCTION

Industrial production and consumption produce unintended side effects that scientists and regulators try to identify and assess. The most important component of such an assessment process is the characterization of risks associated with the activity in question. The normal risk assessment process follows a well-defined protocol of toxicological or epidemiological procedures, which ensure that regulatory or other management actions are based on proven evidence of a potential damage (National Research Council 1983). Until recently, the common wisdom of risk assessors in the field of chemicals have been to distinguish two types of substances: the first groups includes potentially toxic substances that may cause physical damages to human being or the environment above a certain threshold of exposure or intake. Risk management agencies are therefore advised to make sure that the concentration levels would never reach or even surpass these thresholds. With respect to human health, additional safety factors (normally 100-1000 for most risk domains) are applied to adjust for inter-species extrapolation and inter-individual variation. The second class of chemicals is believed to cause harm at any level above zero (stochastic effects). These have been associated with genotoxic effects implying the possibility of irreversible damage to the DNA at an exposure level of a single molecule (one-shot-hypothesis). The regulator has been advised to minimize exposure of people to these stochastic risks (ALARA) and define a level of tolerable risk based on the extrapolation from large to small doses.

This conventional view of toxicity and risk has lately been challenged in the risk assessment and management community. First, the ALARA principle does not specify what "reasonable" means and how much effort needs to be invested in order to reduce risks to a level deemed acceptable. Second, the extrapolation from high to low doses must be done on the basis of a theoretical assumption about the slope and shape of the dose-response function. There are always more than one possibility to draw a regression line from empirical known effects (at high dose) to theoretically modeled low dose effects. Until today there is no community-wide agreement about the shape and slope when it comes to modeling the effects of risk-inducing substances or radiation. In order to avoid these two problems some risk assessors and risk assessment agencies have proposed to use another method, i.e. the *Margin of Exposure Approach* (preferably the BMDL or alternative the T25). This method is based on the definition of a benchmark on a given empirically derived dose-response curve (Dybing et al. 2003; US-EPA 1995; EFSA 2005). The preferred benchmark suggested by the proposers is the point of the dose response curve where 10% of the investigated species show the targeted negative health effect, i.e. the development of a tumor, at the 5% mark representing the upper boundary of the 95% confidence interval. Alternatively the 25% point has also be selected if the 10% mark is impossible or difficult to determine.

The dose that is equivalent to the chosen benchmark is then divided by the actual exposure in humans. The result of this division signals the distance from the 10% level to the actual intake. The authors of the EFSA report suggest that a result of 10,000 and more could be interpreted as a tolerable risk level. A factor of 10,000 means that the concentration of this substance is 10,000 times smaller than the concentration by which at the 5% confidence level 10 percent of the investigated species would develop a tumor. Accordingly, if the 25% benchmark were used, the tolerable level would be divided by an additional 2.5. The same report stresses that the MOE approach does not pre-determine the level of tolerability or acceptability but may help to provide a better comparative foundation for setting priorities. Any substance that has a lower MOE than another substance is potentially more dangerous and would need more attention. Whether 10,000 or any other number should be used to serve as the demarcation line between acceptable and unacceptable risks, is clearly a political decision which needs to be made by risk managers (may be in conjunction with stakeholder involvement). The MOE method itself provides a more reliable base for comparing different substances and facilitating the setting of political standards but does not pre-impose a special tolerability threshold.

However, the picture becomes even more complex if hromesis is also taken into account (Calabrese et al. 1999). Hormesis has been defined as a dose-response relationship in which there is a stimulatory response at low doses, but an inhibiting response at high doses, resulting in a U- or inverted U-shaped dose response (Calabrese and Baldwin 2001). These hormesis effects have been studied for more than two decades (see for example Stebbing 1982; 1998). Toxic agents that are detrimental to human health above certain threshold levels may induce positive effects at a dose that is significantly lower than the NOAEL level. Many recent publications (including those

collected in this volume) provide impressive evidence for the existence of such positive stimulatory effects of low dose exposure. Calabrese and Baldwin report that 19.5% of 1089 samples showed a clear positive hormesis effect, in 80% of the cases such hormesis effect could not be statistically proven (no significant difference to the control group), yet only 0.6% turned out to be false-positive candidates (Calabrese and Baldwin 2001, p. 350). In spite of the evidence for hormesis effects, the topic is still a matter of high controversy among toxicologists.

Until now, regulatory agencies have been reluctant to address this new challenge or adjusted their routines for regulating such substances. All regulatory regimes throughout the world are still based on the traditional risk model: either to define a standard based on thresholds modified by appropriate safety factors or to define tolerable risk levels for stochastic risks caused by chemicals or physical agents (such as radiation). If the hormesis thesis were to be recognized by the scientific community as the new valid paradigm of dose-effect relationships, regulatory systems would need an adjustment and develop new management rules for dealing with the potentially positive effects of low dose exposure. The recent proposal of using a Margin of Exposure approach could be modiefied by applying less stringent distance parameters (varying from 10,000 to 1,000) if hormesis effects are likely to occur at these low exposure levels.

If we turn to the public, the effect of the debate on public opinion so far is confusion. Most people simply demand healthy and safe products and like to act on the assumption ”better safe than sorry” (Lee 1981). This attitude is likely to encourage regulators to err on the safe side and continue to “ignore” potential hormesis effects. At the same time, however, people as consumers have an interest in a large variety of products, low prices and job opportunities. Unless risk information explicitly addresses aspects of potential benefits and social needs, it will not correspond to the expressed and revealed preferences of the people it is supposed to serve.

Based on these considerations about major risk trade-offs in risk management, it is essential to review the ethical implications of hormesis in risk assessment and management. What kind of values should govern the regulation of substances and radiation that may cause positive and negative impacts at the same time (depending on dose and individual variability)? This paper tries to address this question. It is divided into two major sections. Section 1 and 2 will introduce the essentials of ethics and the application of ethical principles to judging the acceptability of risks to humans and the environment. Sections 3 and 4 address the application of these principles to risk management taking into account the hormesis challenge[1].

2. BASIC ETHICAL PRINCIPLES

Should people be allowed to do everything that they are capable of doing? This question is posed in connection with new technologies, hazardous substances, radiation, or human interventions into the natural environment. Intuitively everyone answers this question with a definitive “No“: No way should people be allowed to do everything that they are capable of doing. This also applies to everyday actions. Many options in daily life, from lying to minor deception, from breaking a promise up to going behind a friend’s back, are obviously actions that are seen by all well-intentioned observers as unacceptable. However, it is much more difficult to assess those actions where the valuation is not so obvious. Is it justified to break a promise when keeping the promise could harm many other people?

Actions where there are conflicts between positive and negative consequences or where a judgement could be made one way or the other with equally good justification are especially common in the field of risk evaluation and management. There is hardly anyone who wilfully and without reason exposes people to a health risks, releases toxic pollutants or damages the environment. People who pursue their own selfish goals on the cost and risk of others are obviously acting wrongly and every legislator will sanction this behaviour with the threat of punishment or a penalty. But there is a need for clarification where people bring about a benefit to society with the best intentions and for plausible reasons and, in the process, risk negative impacts on others. In ethics we talk about “conflicting values“ here.

Most decisions involving risks to oneself or others are made for some reason: the actors who make such interventions want to secure good or services to consumers, for example, to ensure long-term jobs and adequate incomes, to incorporate potentially hazardous material for products and services or to use natural reservoirs (sinks) for disposing of waste materials from production and consumption. None of this is done for reasons of brotherly love, but to maintain social interests. Even improving one’s own financial resource is not immoral merely for this reason. The list of human activities that pose risks onto others perpetrated for existential or economic reasons could be carried on into infinity. Human existence is bound to taking opportunities and risks.

Therefore, to be able to make a sensible judgement of the balance between necessary improvements of the present status of society and the risks to human health and environmental quality posed by these activities, the range of products and services has to be systematically compared to the losses that are inflicted on human health and the environment. If important goods have to be appreciated when weighing the pros and cons of human activities. criteria are needed that can be used as yardsticks. Who can and may draw up such criteria, according to which standards should the risk inducing activities be assessed and how can the various evaluative options for action be compared with each other for each criterion?

Taking risks always involves two major components: an assessment of what we can expect from the activity and an evaluation of how desirable these expectations are. The first component addresses the risk and benefit *assessment* side of the risk analysis. The second component addresses the *societal evaluation* of these expected consequences. Whereas the estimate of consequences broadly falls in the domain of scientific research and expertise, with uncertainties and ambiguities in particular having to be taken into account (IRGC 2005), the question about the foundations for evaluating various options for action and about drawing up standards guiding action is a central function of ethics. Ethics can provide an answer to the question posed at the

beginning ("Should people be allowed to do everything that they are capable of doing?") in a consistent and transparent manner.

3. ETHICAL FOUNDATIONS OF RISK MANAGEMENT

3.1 Overview of ethical approaches to risk

Answering the question about the finding the right balance between benefits and risks lies within the field of practical philosophy, ethics. Following the usual view in philosophy, ethics describes the theory of the justification of normative statements, i.e. those that guide action (Gethmann, 1991; Mittelstraß, 1992; Nida-Rümelin, 1996a; Revermann, 1998). A system of normative statements is called "morals". Ethical judgements therefore refer to the justifiability of moral instructions for action that may vary from individual to individual and from culture to culture (Ott, 1999).

Basically, humans are purpose-oriented and self-determined beings who act not only instinctively, but also with foresight, and are subject to the moral standards to carry out only those actions that they can classify as good and justifiable (Honnefelder, 1993). Obviously, not all people act according to the standards that they themselves see as necessary, but they are capable of doing so. In this context it is possible for people to act morally because, on the one hand, they are capable of distinguishing between moral and immoral action and, on the other, are largely free to choose between different options for action.

Whether pursuing a particular instruction for action should be considered as moral or immoral is based on whether the action concerned can be felt and justified to be "reasonable" in a particular situation. Standards that cross over situations and that demand universal applicability are referred to as *principles* here. Conflicts may arise between competing standards (in a specific situation), as well as between competing principles, the solution of which, in turn, needs justification (Szejnwald-Brown et al., 1993). Providing yardsticks for such justification or examining moral systems with respect to their justifiability is one of the key tasks of practical ethics (Gethmann, 1998).

In ethics a distinction is made between descriptive (experienced morality) and prescriptive approaches, i.e. justifiable principles of individual and collective behaviour (Frankena, 1963; Hansen, 1995). Furthermore, ethical norms can be a applied to the personal lifestyle ("good life") and to collective actions (normative guidelines) (Galert, 1998; Ott, 1999). Within normative ethics a distinction is made between deontological and teleological approaches when justifying normative judgments (Höffe, 1987). Deontological approaches are principles and standards of behaviour that apply to the behaviour itself on the basis of an external valuation criterion. It is not the consequences of an action that are the yardstick of the valuation; rather it is adhering to inherent yardsticks that can be used against the action itself. Such external yardsticks of valuation are derived from religion, nature, intuition or common sense, depending on the basic philosophical direction. Thus, protection of the biosphere can be seen as a divine order to protect creation (Rock, 1980; Schmitz, 1985), as an innate tendency for the emotional attachment of people to an environment with biodiversity (Wilson, 1984), as a directly understandable source of inspiration and joy (Ehrenfeld, 1993) or as an educational means of practising responsibility and maintaining social stability (Gowdy, 1997).

By contrast, teleological approaches refer to the consequences of action. Here, too, external standards of valuation are needed since the ethical quality of the consequences of action also have to be evaluated against a yardstick of some kind. With the most utilitarian approaches (a subset of the teleological approaches) this yardstick is defined as an increase in individual or social benefit. In other schools of ethics, intuition (can the consequence still be desirable?) or the aspect of reciprocity (the so-called "Golden Rule" "do as you would be done by") play a key role.

In the approaches based on logical reasoning (especially in Kant), the yardstick is derived from the logic of the ability to generalise or universalise. Kant himself is in the tradition of deontological approaches ("Good will is not good as a result of what it does or achieves, but just as a result of the intention"). According to Kant every principle that, if followed generally, makes it impossible for a "good life" to be conducted is ethically impermissible. In this connection, it is not the desirability of the consequences that captures Kant's mind, but the logical inconsistency that results from the fact that the conditions of the actions of individuals would be undermined if everyone were to act according to the same maxims (Höffe, 1992).

A number of contemporary ethicists have taken up Kant's generalisation formula, but do not judge the maxims according to their internal contradictions; rather they judge them according to the desirability of the consequences to be feared from the generalisation (Jonas, 1979 or Zimmerli, 1993). These approaches can be defined as a middle course between deontological and teleological forms of justification.

In addition to deontological and teleological approaches there is also the simple solution of consensual ethics, which, however, comprises more than just actually experienced morality. Consensual ethics presupposes the explicit agreement of the people involved in an action. Everything is allowed provided that all affected (for whatever reason) voluntarily agree. In sexual ethics at the moment a change from deontological ethics to a consensual moral code can be seen.

The comparison of the basic justification paths for normative moral systems already clearly shows that professional ethicists cannot create any standards or designate any as clearly right, even if they play a role in people's actual lives. Much rather it is the prime task of ethics to ensure on the basis of generally recognised principles (for example, human rights) that all associated standards and behaviour regulations do not contradict each other or a higher order principle.

Above and beyond this, ethics can identify possible solutions that may occur with a conflict between standards and principles of equal standing. Ethics may also reveal interconnections of justification that have proved themselves as examination criteria for moral action in the course of their disciplinary history. Finally, many ethicists see their task as providing methods and procedures primarily of an intellectual

nature by means of which the compatibility or incompatibility of standards within the framework of one or more moral systems can be completed. Unlike the law, the wealth of standards of ethics is not bound to codified rules that can be used as a basis for such compatibility examinations. Every normative discussion therefore starts with the general issues that are needed in order to allow individuals a "good life" and, at the same time, to give validity to the principles required to regulate the community life built on common good. But how can generally binding and intersubjectively valid criteria be made for the valuation of "the common good"?

3.2 The problem of ultimate justification

In modern pluralistic societies it is increasingly difficult for individuals and groups of society to draw up or recognise collectively binding principles that are perceived by all equally as justifiable and as self-obliging (Hartwich and Wewer, 1991; Zilleßen, 1993). The variety of lifestyle options and subjectification of meaning (individualisation) are accompanying features of modernisation. With increasing technical and organisational means of shaping the future, the range of behaviour options available to people also expands. With the increasing plurality of lifestyles, group-specific rationalities emerge that create their own worldviews and moral standards, which demand a binding nature and validity only within a social group or subculture. The fewer cross-society guiding principles or behaviour orientations are available, the more difficult is the process of agreement on collectively binding orientations for action. However, these are vital for the maintenance of economic cooperation, for the protection of the natural foundations of life and for the maintenance of cohesion in a society. No society can exist without the binding specification of minimum canons of principles and standards.

But how can agreement be reached on such collectively binding principles and standards? What criteria can be used to judge standards? The answers to this question depend on whether the primary principles, in other words the starting point of all moral systems, or secondary principles or standards, i.e. follow-on standards that can be derived from the primary principles, are subjected to an ethical examination. *Primary principles can be categorical or compensatory* (capable of being compensated). Categorical principles are those that must not be infringed under any circumstances, even if other principles would be infringed as a result. The human right to the integrity of life could be named here as an example. Compensatory principles are those where temporary or partial infringement is acceptable, provided that as a result the infringement of a principle of equal or higher ranking is avoided or can be avoided. In this way certain freedom rights can be restricted in times of emergency. In the literature on ethical rules, one can find more complex and sophisticated classifications of normative rules. For our purpose to provide a simple and pragmatic framework, the distinction in four categories (principles and standards; categorical and compensatory) may suffice..

But how can primary principles be justified as equally valid for all people? Although many philosophers have made proposals here, there is a broad consensus today that neither philosophy nor any other human facility is capable of stating binding meta-criteria without any doubt and for all people, according to which such primary principles should be derived or examined (Mittelstraß, 1984). A final justification of normative judgements cannot be achieved by logical means either, since all attempts of this kind automatically end either in a logical circle, in an unending regression (vicious cycle) or in a termination of the procedure and none of these alternatives is a satisfactory solution for final justification (Albert, 1991).

The problem of not being able to derive finally valid principles definitively, however, seems to be less serious than would appear at first glance. Because, regardless of whether the basic axioms of moral rules are taken from intuition, observations of nature, religion, tradition reasoning or common sense, they have broadly similar contents. Thus, there is broad consensus that each human individual has a right to life, that human freedom is a high-value good and that social justice should be aimed at. But there are obviously many different opinions about what these principles mean in detail and how they should be implemented. In spite of this plurality, however, discerning and well-intentioned observers can usually quickly agree, whether one of the basic principles has clearly been infringed. It is more difficult to decide whether they have clearly been fulfilled or whether the behaviour to be judged should clearly be assigned to one or several principles. Since there is no finally binding body in a secular society that can specify primary principles or standards *ex cathedra*, in this case consensus among equally defendable standards or principles can be used (or pragmatically under certain conditions also majority decisions). Ethical considerations are still useful in this case as they allow the test of generalisation and the enhancement of awareness raising capabilities. In particular, they help to reveal the implications of such primary principles and standards.

Provided that primary principles are not concerned (such as human rights), the ethical discussion largely consists of examining the compatibility of each of the available standards and options for action with the proposed primary principles. In this connection, the main concerns are a lack of contradictions (consistency), logical consistency (deductive validity), coherence (agreement with other principles that have been recognised as correct) and other, broadly logical criteria (Gethmann, 1998). As the result of such an examination it is entirely possible to reach completely different conclusions that all correspond to the laws of logic and thus justify new plurality.

In order to reach binding statements or valuations here the evaluator can either conduct a discussion in his "mind" and let the arguments for various standards compete with each other (rather like a platonic dialogue) or conduct a real discussion with the people affected by the action. In both cases the main concern is to use the consensually agreed primary principles to derive secondary principles of general action and standards of specific action that should be preferred over alternatives that can be equally justified. A plurality of solutions should be expected especially because most of the concrete options for action comprise only a gradual fulfilment and infringement of primary principles and therefore also include conflicting values. For value conflicts at the same level of abstraction there are, by definition, no clear rules for solution. There are therefore frequently conflicts between conserving life through economic development and destroying life through hazardous materials. Since the principle of conserving

life can be used for both options a conflict is unavoidable in this case. To solve the conflicts ethical considerations, such as the avoidance of extremes, staggering priorities over time or the search for third solutions can help without, however, being able to convincingly solve this conflict in principle to the same degree for all (Szejnwald-Brown et al., 1993).

These considerations lead to some important conclusions for the matter of the application of ethical principles to the issue of human action with regard to risks to human health and the natural environment. First of all, it contradicts the way ethics sees itself to develop ethics of its own for different action contexts. Just as there can be no different rules for the logic of deduction and induction in nomological science, depending on which object is concerned, it does not make any sense to postulate an independent set of ethics for risk management concerning effects on human health or the environment (Galert, 1998). Justifications for principles and moral systems have to satisfy universal validity (Nida-Rümelin, 1996b).

Therefore, it is not helpful to call for a special moral system for evaluating risks since this – like every other moral system – has to be traceable to primary principles. Instead, it makes sense to specify the generally valid principles that are also relevant with regard to the issue of how to deal with risks and benefits of human activities. At the same time standards should be should be derived from these principles that provide concrete guidelines of how to balance risks and benefits.

3.3 Categorical versus compensatory principles and standards

With regard to risk and benefits of human activities, different goods have to be weighed up against each other. There is no magic formula available indicating how much risk can be traded for how much valuable commodities. Humans alone are responsible for the resolution of conflicts between competing objectives. Appreciation and negotiation processes are therefore at the core of the considerations about ethical principles and standards of risk acceptability.

But this does not mean that there is no room for categorical judgements along the lines of "this or that absolutely must be prohibited" in the matter of risk evaluation. It follows on from the basic principle of conserving human life that all human interventions that threaten the ability of the human race as a whole, or a significant number of individuals alive today or in the future, to exist should be categorically prohibited. This refers to risks that threaten the systemic functions of the biosphere. Such threats are one of the guiding principles that must not be exceeded under any circumstances, even if this excess were to be associated with high benefits. In the language of ethics this is a categorical principle, in the language of economics a good that is not capable of being traded.

A second non-negotiable categorical norm is the protection of individual human lives unless other lives are jeopardised There are many exceptions to this categorical law. It is, for example, morally not justified to kill one person and use his or her organs to save two other persons. Without going into much detail here, imposing risks which are very likely to kill other individuals or to seriously damage their health are not justified regardless what economic benefit is associated with these risks. However, below the threshold of serious risks, some imposition of risks onto others (ideally with their informed consent) is legitimate if these risks are balanced with major benefits to society (Shrader-Frechette 1991). In this case risk to life can be compensated with other goods. In the past, a number of authors have tried to specify the minimum requirements for acceptable risk levels (from which on compensation is legitimate). These so-called "safe minimum standards" specify thresholds for the measurement scale of risks (between 0 and 1) that may not be exceeded even if there is a prospect of great benefits (Randall, 1988; Randall and Farmer, 1995).

For most risks caused by chemical substances or radiation one can assume that compensatory rules apply. If indeed a risk would exceed the tolerable risk level set by societal consensus a release of such a chemical or physical risk would not be permitted. In all other cases the risk of being harmed by a substance or a release of radiation needs to be compared with the benefit of the activity that is associated with the risk in question. In order to evaluate partial infringements of compensatory principles or standard society needs rules for decision-making that facilitate the balancing process necessary to resolve compensatory conflicts. In the current debate about rules for risk management it is mainly teleological valuation methods that are proposed (Hubig, 1993; Ott, 1993).

These methods are aimed at:

> estimating the possible consequences of various options for action at all dimensions relevant to potentially affected people,
>
> recording the infringements or fulfilments of these expected consequences in the light of the existing standards and principles and
>
> then weighting them according to an internal key so that they can be weighed up in a balanced way.

On the positive side of the equation there are the economic benefits of a risk-inducing activity and the cultural values created by its application, for example in the form of income, health enhancement or an aesthetically attractive landscape (parks, ornamental gardens, etc.); on the negative side there are threats to human health, the natural environment or the violation of aesthetic, cultural or religious attributes associated with the respective risk taking.

In risk-benefit assessment there are frequently related categories on both sides of the equation: With the same or similar categories on the credit and debit side of the balance sheet the decision is easy when there is one option that performs better or worse than all the other options for all categories. Such a *dominant* (the best for all categories) or *sub-dominant option* (the worst for all categories) is, however, rare in reality. If we disregard the dominant or sub-dominant solutions, an appreciation between options that violate or fulfil compensatory standards and principles depends on two preconditions: best possible knowledge of the consequences (what happens if I choose option A instead of option B?) and a transparent, consistent rationale for weighing up these consequences as part of a legitimate political decision process (are the foreseeable consequences of A more desirable

or bearable than the consequences of option B?) (Akademie der Wissenschaften, 1992).

3.4 Knowledge and values as a basis for risk assessment and management

In order to conduct such an informed balance one needs, first of all, adequate knowledge about the likely consequences in order to reveal the systemic connections between a human activity and its impacts on all dimensions that humans value (Wolters, 1995). This requires interdisciplinary research and cooperation. The task of toxicology in this multidisciplinary exercise, for example, is to show the consequences of using a specific substance on human health and ecological systems. The economic disciplines provide a benefit-oriented valuation of the application of this substance in different products and demonstrate the impacts for economy and well-being of all affected individuals. Cultural and social sciences investigate the feedback effects between this application, social development and cultural self-perception. They illustrate the dynamic interactions between exposure, socio-cultural lifestyles and control options. Interdisciplinary, problem-oriented and system-related research is needed to contribute to forming a basic stock of findings and insights about functional links in the relationship between risk-inducing human activities and their consequences on human health and the environment (WBGU, 2000).

But knowledge alone does not suffice. In order to be able to act effectively and efficiently while observing ethical principles, it is necessary to shape the *evaluation process* between the various options for action *according to rational criteria* (Gethmann, 1998). To do this it is first of all necessary to identify the dimensions that should be used for the evaluation of risks. The discussion about the value dimensions to be used as a basis for evaluation is one of the most popular subjects within environmental ethics. To apply these criteria in risk evaluation and to combine the knowledge aspects about expected consequences of different behavioural options with the ethical principles is the task of what we have called risk governance (IRGC 2005). Within risk governance the main criteria are:

Effectiveness: Does the activity and/or the risk management option achieve the desired effect?

Efficiency: Does the activity and/or the risk management option achieve the desired effect with the least resource consumption?

Minimisation of external side effects: Does the activity and/or the risk management option infringe on other valuable goods, benefits or services such as competitiveness, public health, environmental quality, social cohesion, etc.? Does it impair the efficiency and acceptance of the governance system itself?

Sustainability: Does the activity and/or the risk management option contribute to the overall goal of sustainability? Does it assist in sustaining vital ecological functions, economic prosperity and social cohesion?

Fairness: Does the activity and/or the risk management option burden the subjects of regulation in a fair and equitable manner?

Political and legal implementability: Is the activity and/or the risk management option compatible with legal requirements and political programmes?

Public acceptance: Will the activity and/or the risk management option be accepted by those individuals who are affected by it? Are there cultural preferences or symbolic connotations that have a strong influence on how the risks are perceived?

Measuring risk-inducing activities or risk reducing management options against these criteria may create conflicting messages and results. Many measures that prove to be effective may turn out to be inefficient or unfair to those who will be burdened. Other measures may be sustainable but not accepted by the public or important stakeholders. There are many excellent guidance documents available that demonstrate how to handle painful risk trade-offs and how to employ decision analytic tools for dealing with conflicting evidence and values (c.f. Viscusi 1994; Wiener 1998; van der Sluijs et al. 2003; Goodwin and Wright 2004). The following section will present a framework for applying these principles to risk management with special emphasis on hormesis. The main line of argument is that risk management requires an analytic-deliberative approach for dealing effectively and prudently with complex risks.

4. AN ANALYTIC-DELIBERATIVE APPROACH TO EVALUATING COMPLEX RISKS

4.1 Combining ethical evaluation and risk management

Assessing potential consequences of human interventions and evaluating their desirability on the basis of subsequent knowledge and transparent valuation criteria are two of the central tasks of a risk governance process. However, the plural values of an heterogeneous public and people's preferences have to be incorporated in this process. But how can this be done given the wealth of competing values and preferences? Should we simply accept the results of opinion polls as the basis for making political decisions? Can we rely on risk perception results to judge the seriousness of pending risks? Or should we place all our faith in professional risk assessment and management?

If we turn to professional help to deal with plural value input, economic theory might provide us an answer to this problem: If environmental goods are made individual and suitable for the market by means of property rights, the price that forms on the market ensures an appropriate valuation of the good. Every user of this good can then weigh up whether he is willing to pay the price or would rather not use the good. With many goods that could pose a health threat to humans, however, this valuation has to be made by collective action because public health good is a collective good that cannot be governed by individual action. In this case a process is needed that safeguards the collective rationale in valuation and justifies it to the collective. However, this valuation cannot be determined with the help of survey results. Although surveys are needed to be able to estimate the breadth of preferences and people's willingness to pay, they are insufficient for a derivation of concrete decision-making criteria and

yardsticks for evaluating the tolerability of risks to human health and the environment (Shrader-Frechette 1991).

Firstly, the individual values are so widely scattered that there is little sense in finding an average value here.

Secondly, the preferences expressed in surveys change so much within short time whereas ethical valuations have to be valid for a long time.

Thirdly, individual preferences are frequently based on flawed knowledge or ad hoc assumptions both of which should not be decisive according to rational considerations.

What is needed, therefore, is a gradual process of assigning trade-offs in which existing empirical values are put into a coherent and logically consistent form.

In political science and sociological literature reference is mostly made to three strategies of incorporating social values and preferences in rational decision-making processes (Renn, 1997). Firstly, a reference to social preferences is viewed solely as a question of legitimate procedure (Luhmann, 1983; Vollmer, 1996). The decision is made on the basis of formal decision making process(such as majority voting). If all the rules have been kept a decision is binding, regardless of whether the subject matter of the decision can be justified or whether the people affected by the decision can understand the justification. In this version, social consensus has to be found only about the structure of the procedures; the only people who are then involved in the decisions are those who are explicitly legitimated to do so within the framework of the procedure decided upon.

The second strategy is to, rely on the minimum consensuses that have developed in the political opinion-forming process (muddling through) (Lindbloom, 1959, 1965). In this process, only those decisions that cause the least resistance in society are considered to be legitimate. In this version of social pluralism groups in society have an influence on the process of the formation of will and decision-making to the extent that they provide proposals capable of being absorbed, i. e. adapted to the processing style of the political system, and that they mobilise public pressure. The proposal that then establishes itself in politics is the one that stands up best in the competition of proposals, i.e. the one that entails the fewest losses of support for political decision-makers by interest groups.

The third strategy is based on the discussion between the groups involved (Habermas, 1971, 1991; Renn 2004). In the communicative exchange among the people involved in the discussion a form of communicative rationality that everyone can understand evolves that can serve as a justification for collectively binding decisions. At the same time, discursive methods claim to more appropriately reflect the holistic nature of human beings and also to provide fair access to designing and selecting solutions to problems. In principle the justification of standards relevant to decisions is linked to two conditions: the agreement of all involved and substantial justification of the statements made in the discussion (Habermas, 1981).

All three strategies of political control are represented in modern societies to a different extent. Legitimisation conflicts mostly arise when the three version are realised in their pure form. Merely formally adhering to decision-making procedures without a justification of content encounters a lack of understanding and rejection among the groups affected especially when they have to endure negative side effects or risks. Then acceptance is refused. If, however, we pursue the opposite path of least resistance and base ourselves on the route of muddling through we may be certain of the support of the influential groups, but, as in the first case, the disadvantaged groups will gradually withdraw their acceptance because of insufficient justification of the decision. At the same time, antipathy to politics without a line or guidance is growing, even the affected population. The consequence is political apathy.

The third strategy of discursive control faces problems too. Although in an ideal situation it is suitable for providing transparent justifications for the decision-making methods and the decision- itself, in real cases the conditions of ideal discourse can rarely be adhered to (Wellmer, 1992). Frequently, discussions among strategically operating players lead to a paralysis of practical politics by forcing endless marathon meetings with vast quantities of points of order and peripheral contributions to the discussion. The "dictatorship of endurance" (Weinrich, 1972) ultimately determines which justifications are accepted by the participants. The public becomes uncertain and disappointed by such discussions that begin with major claims and end with trivial findings.

In brief: none of the three ways out of the control dilemma can convince on its own; as so often in politics, everything depends on the right mixture. What should a mixture of the three elements (due process, pluralistic muddling through and discourse) look like so that a maximum degree of rationality can come about on the basis of social value priorities?

A report by the American National Academy of Sciences on the subject of "Understanding environmental risks" (Stern and Fineberg, 1996) comes to the conclusion that scientifically valid and ethically justified procedure for the collective valuation of options for risk handling can only be realised within the context of –what the authors coin— an analytic-deliberative process. *Analytic* means that the best scientific findings about the possible consequences and conditions of collective action are incorporated in the negotiations; *deliberative* means that rationally and ethically transparent criteria for making trade-offs are used and documented externally. Moreover, the authors consider fair participation by all groups concerned is necessary to ensure that the different moral systems that can legitimately exist alongside each other should also be incorporated in the process.

To illustrate the concept of analytic-.deliberative decision making consider a set of alternative *options* or *choices*, from which follow *consequences* (see basic overview in Dodgson et al. 2000). The relationship between the choice made, and the consequences that follow from this choice, may be straightforward or *complex*. The science supporting risk management is often complicated, across many disciplines of science and engineering, and also involving human institutions and economic interactions. Because of limitations in scientific understanding and predictive capabilities, the consequences following a choice are normally *uncertain*. Finally, different individuals and groups with-

in society may not agree on how to evaluate the consequences – which may involve a detailed characterization of what happens in ecological, economic, and human health terms. We shall describe consequences as *ambiguous* when there is this difficulty in getting agreement on how to interpret and evaluate them.

Risk management inherently involve these difficulties of *complexity, uncertainty, and ambiguity* (Klinke and Renn 2002). In some situations where there is lots of experience, these difficulties may be minimal. But in other situations these difficulties may constitute major impediments to the decision making process. To understand how analysis and deliberation interact in an iterative process following the NRC 1996 report, one must consider how these three areas of potential difficulty can be addressed. It is useful to separate questions of evidence with respect to the likelihood, magnitude of consequences and related characteristics (which can involve *complexity* and *uncertainty*) from socio-political evaluation of the consequences (i.e. *ambiguity*). For each of the three areas there are analytical tools that can be helpful in identifying, characterizing and quantifying cause-effect relationships. The integration of these tools of risk governance into a consistent procedure will be discussed in the next subsections.

4.2 Analytic-deliberative processes: Towards a procedural integration

The possibility to reach closure on evaluating risks to human health or the environment rests on two conditions: first, all participants need to achieve closure on the *underlying goal* (often legally prescribed such as prevention of health detriments or guarantee of an undisturbed environmental quality, for example purity laws for drinking water); secondly, they need to agree with the implications derived from the *present state of knowledge* (whether and to what degree the identified hazard impacts the desired goal). Dissent can result from conflicting values as well as conflicting evidence. It is crucial in risk management to investigate both sides of the coin: the values that govern the selection of the goal and the evidence that governs the selection of cause-effect claims.

Separating the science issues of what will happen from the value issues of how to make appropriate tradeoffs between ecological, economic, and human health goals can become very cumbersome. The separation of facts and values in decision making is difficult to accomplish in practical decision situations, since what is regarded as facts includes a preference dependent process of cognitive framing (Tversky and Kahneman 1981) and what is regarded as value includes a prior knowledge about the factual implications of different value preferences (Fischhoff 1975). Furthermore, there are serious objections against a clear cut division from a sociological view on science and knowledge generation (Jasanoff 1996). Particularly when calculating risk estimates, value-based conventions may enter the assessment process. For example, conservative assumptions may be built into the assessment process, so that some adverse effects (such as human cancer from pesticide exposure) are much less likely to be underestimated than overestimated (National Research Council 1983). Similarly the decision to include or exclude potential hormesis effects may alter the final judgment about acceptability of a given exposure. At the same time, ignoring major sources of uncertainty can evoke a sense of security and overconfidence that is not justified from the quality or extent of the data base (Einhorn and Hogarth 1978). Perceptions and world views may be very important, and difficult to sort out from matters of science, especially with large uncertainties about the risks in question.

A combination of analytic and deliberative processes can help explore these differences of opinions relating to complexity, uncertainty, and ambiguity in order to examine the appropriate basis for a decision before the decision is made. Most risk management agencies go through an elaborate assessment process and provide opportunities for public review and comment. Many controversial risk decisions become the focus of large analytical efforts, in which mathematical models are used to predict the environmental, economic, and health consequences of different management alternatives. Analysis should be seen as an indispensable complement to deliberative processes, regardless whether this analysis is sophisticated or not. Even simple questions need analytic input for making prudent decisions, especially in situations where there is controversy arising from complexity, uncertainty, and ambiguity.

4.3 Conducting deliberations on risks issues

In the course of practical risk management different conflicts arise in deliberative settings that have to be dealt with in different ways. The main conflicts occur at the process level (how should the negotiations be conducted?), on the cognitive level (what is factually correct?), the interest level (what benefits me?), the value level (what is needed for a "good" life?) and the normative level (what can I expect of all involved?). These different conflict levels are addressed in this subsection.

First of all, negotiations begin by specifying the method that structures the dialogue and the rights and duties of all participants. It is the task of the chairman or organiser to present and justify the implicit rules of the talks and negotiations. Above and beyond this, the participants have to specify joint rules for decisions, the agenda, the role of the chairman, the order of hearings, etc. This should always be done according to the consensus principle. All partners in the negotiations have to be able to agree to the method. If no agreement is reached here the negotiations have to be interrupted or reorganised.

Once the negotiation method has been determined and, in a first stage, the values, standards and objectives needed for judgement have been agreed jointly, then follows the *exchange of arguments and counter arguments*. In accordance with decision theory, four stages of validation occur:

In a first stage, the values and standards accepted by the participants are translated into criteria and then into indicators (measurement instructions). This translation needs the consensual agreement of all participants. Experts are asked to assess the available options with regard to each indicator according to the best of their knowledge (factual correctness). In this context it makes more sense to specify a joint methodological procedure or a consensus about the experts to be questioned than to give each group the freedom to have the indicators answered by their own experts. Often many potential consequences

remain disputed as a result of this process, especially if they are uncertain. However, the bandwidth of possible opinions is more or less restricted depending on the level of certainty and clarity associated with the issue in question. Consensus on dissent is also of help here in separating contentious factual claims from undisputed ones and thus promotes further discussion.

In a second stage, all participating parties are required to interpret bandwidths of impacts to be expected for each criterion. *Interpretation* means linking factual statements with values and interests to form a balanced overall judgement (conflicts of interests and values). This judgement can and should be made separately for each indicator. In this way each of the chains of causes for judgements can be understood better and criticised in the course of the negotiations. For example, the question of trustworthiness of the respective risk management agencies may play an important role in the interpretation of an expected risk value. Then it is the duty of the participating parties to scrutinise the previous performance of the authority concerned and propose institutional changes where appropriate.

Third stage: Even if there were a joint assessment and interpretation for every indicator, this would by no means signify that agreement is at hand. Much rather, the participants' different judgements about decision-making options may be a result of different *value weightings* for the indicators that are used as a basis for the values and standards. For example, a committed environmentalist may give much more weight to the indicator for conservation than to the indicator of efficiency. In the literature on game theory this conflict is considered to be insoluble unless one of the participants can persuade the other to change his preference by means of compensation payments (for example, in the form of special benefits), transfer services (for example, in the form of a special service) or swap transactions (do, ut des). In reality, however, it can be seen that participants in negotiations are definitely open to the arguments of the other participants (i.e. they may renounce their first preference) if the loss of benefit is still tolerable for them and, at the same time, the proposed solution is considered to be "conducive to the common good", i.e. is seen as socially desirable in public perception. If no consensus is reached, a compromise solution can and should be reached, in which a "fair" distribution of burdens and profits is accomplished.

Fourth stage: When weighing up options for action *formal methods of balancing assessment* can be used. Of these methods, the cost-benefit analysis and the multi-attribute or multi-criteria decision have proved their worth. The first method is largely based on the approach of revealed "preferences", i.e. on people's preferences shown in the past expressed in relative prices, the second on the approach of "expressed preferences", i.e. the explicit indication of relative weightings between the various cost and benefit dimensions (Fischhoff et al., 1982). But both methods are only aids in weighing up and cannot replace an ethical reflection of the advantages and disadvantages.

Normative conflicts pose special problems because different evaluative criteria can always be classified as equally justifiable or unjustifiable as explained in section 1 of this paper. For this reason, most ethicists assume that different types and schools of ethical justification can claim parallel validity, it therefore remains up to the groups involved to choose the type of ethically legitimate justification that they want to use (Ropohl, 1991; Renn, 2004). Nevertheless, the limits of particular justifications are trespassed wherever primary principles accepted by all are infringed (such as human rights). Otherwise, standards should be classed as legitimate if they can be defended within the framework of ethical reasoning and if they do not contradict universal standards that are seen as binding for all. In this process conflicts can and will arise, e.g. that legitimate derivations of standards from the perspective of Group A contradict the equally legitimate derivations of Group B (Shrader-Frechette, 1988). In order to reach a jointly supported selection of standards either a portfolio of standards that can claim parallel validity should be drawn up or compensation solutions will have to be created in which one party compensates the other for giving up its legitimate options for action in favour of a common option.

When choosing possible options for action or standards, options that infringe *categorical principles*, for example endangering human lives with a high probability and thus exceeding the limits of tolerability. At the same time, all sub-dominant options have to be excluded. Frequently sub-dominant solutions, i.e. those that perform worse than all other options with regard to all criteria at least in the long term, are so attractive because they promise benefits in the short term although they entail certain losses in the long term, even if high interest rates are assumed. Often people or groups have no choice other than to choose the sub-dominant solution because all other options are closed to them due to a lack of resources. If large numbers of groups or even peoples act in this way, global risks become unmanageable (Beck, 1996). To avoid these risks intermediate financing or compensation by third parties should be considered.

5. APPLICATION TO HORMESIS

If one assumes that the hormesis hypothesis is correct and sufficient evidence has been collected to verify its basic claims. a thorough revision of the present paradigms in regulatory philosophy and actions is necessary. The minimization principle on which most of the traditional regulations rests would be in need of either replacement or refinement. If public policy is meant to improve public health and not only to prevent negative effects, there would be a necessity to seek exposure to small doses or at least to ensure that such an exposure is not prohibited by the minimization principle. In the case of toxic substances with a clear NOAEL, only little changes in the regulatory system are required. Individuals may then be advised to seek exposure rather than avoid it as long as the NOAEL threshold is not reached.

Risk management becomes more difficult and controversial if hormesis is applied to stochastic risk agents. Most dose-response models assume a finite probability for developing a detrimental health effect (most often carcinogenic and/or mutagenic effect) as a result of any exposure above zero. These stochastic effects are due to the possibility of irreversible damage to the DNA at an exposure level of a single molecule. If the hormesis hypothesis is applied to carcinogenic substances or radiation, the probability for a tumor inhibition may outweigh the probability of a tumor induction. Under these circum-

stances the situation might occur that a single individual may develop a tumor as a result of an exposure to a very small dose of a carcinogen, while the majority of people may experience positive inhibitory effects. Similar dilemmas can also occur with simple toxic substances if individuals vary in their sensibility towards the beneficial range of exposure in which the positive effects are observed. One individual may experience the positive effects at a different dose range compared to another more sensitive individual. How should a regulator evaluate such a situation? Is it justified to endorse exposure to small concentrations of a proven carcinogenic or toxic substance if there is a chance that a small number of people will probably be negatively affected while the majority enjoys the potential benefits? This question raises equity concerns and leads to difficult policy dilemmas.

The popular question "how safe is safe enough?" would not only need the addition of "how safe is fair enough" but also "what degree of safety implies living less safe than possible". The paradigm of minimization would need to be replaced by a new optimality rule that allows for beneficial effects of low dose exposure. Instruments for reaching this new paradigm are not yet in place and would require more deliberation and policy studies.

What would the be the ethical implications for risk management if faced with such dilemmas?

First, ignoring hormesis would be immoral as the principle of assigning trade-offs between comparable goods require that all (known) benefits and risks are included in the balancing procedure. This is also true even if the effects are still uncertain yet plausible.

Secondly, the juxtaposition of positive and negative impacts of a given risk (and the respective risk reduction measures) is central to finding a morally justified policy. The newly proposed Margin of Exposure Approach may be a good approximation to structuring such a balance sheet. The factor between exposure level by which 10% of the test animals develop a tumor, and the actual exposure to humans provides a good indicator of the level of protection that society would like to impose on risk-causing activities. Similarly one could calculate a hormesis factor based on the maximum beneficial effect to public health in relation to the 10% exposure. One would expect either a logistical or a sinus function starting with the 10% level and then reducing the exposure level step by step until it reaches the proposed minimum divisor of 10,000. Such a juxtaposition of protection level and hormesis level could assist risk managers to look for the appropriate trade-offs.

Third, the assignment of trade-offs between the potential detrimental effect of a given stochastic risk and the beneficial effect of hormesis as indicated by the MoE approach needs to include equity considerations, basic human rights, and values pertaining to social cohesion, integration and peaceful conflict resolution. It is not sufficient to count the people most likely to receive a benefit and weight them against those that have a higher probability of being harmed. The complexity of finding the appropriate criteria for such a comparative review requires a discursive approach to decision making as explained in earlier sections of this paper.

Fourth, the discourse to find the appropriate trade-offs need to include those who would benefit from the activity (economic actors as well as those who are most likely to benefit from hormesis) und those who would most likely suffer from an exposure (most vulnerable groups). The main objective of such a participatory discourse is the creation of an informed consensus. All parties need to learn what is known about the potential impacts of a risk and the choices among the risk reduction options. They need to be informed about the remaining uncertainties and ambiguities associated with each impact. Based on this common knowledge they can start deliberating about the relative weights to be given to each impact category. The tools of decision analysis can assist the participants to adhere to formal criteria such as consistency, coherence and proportionality. Yet the trade-offs themselves are not pre-determined and cannot be pre-determined according to our analysis on ethics and decision making. The discourse is the place where the various arguments are exchanged and a consensus or at least a compromise might arise from the exchange of ethically informed arguments.

Fifth, the outcome of such a discourse may not provide a general rule for treating stochastic risks with known hormesis effects. It may be specific to different contexts (for example: voluntary exposure versus non voluntary), to different agents (chemicals in food versus chemicals in consumer products) or to different target groups (general population versus special vulnerable groups). What should be expected from such a discourse is not unity but convincing justification for each case.

A regulatory system that has incorporated such a discourse for trading off positive hormesis effects against negative stochastic risks is not in sight. Until now any consideration about hormesis has been excluded from the regulatory systems worldwide, partially because the evidence is still regarded as insufficient to trigger any regulatory action, partially because agencies fear the complexity and ambiguity when faced with positive and negative impacts of the same effect. But with more and more evidence coming in it will be difficult for agencies to ignore the positive effects and, as mentioned before, it would become immoral to ignore such evidence that could help people to improve their health status. Therefore, it is ethically mandated that provisions for including hormesis effects in risk management are introduced and implemented. This being said, the inclusion would not automatically lead to lower standards or a lax regulatory practice. It may be the result of an analytic-deliberative process that the discourse participants agree to place more weight on preventing stochastic genotoxic risks than on positive immunization effects caused by hormesis. If the arguments for both sides are truly considered and weighted against each other the ethical rule of balancing is met.

6. CONCLUSIONS

The objective of this paper was to address and discuss the use of ethical principles and decision analytic tools for standard setting procedures in risk management with special emphasis on hormesis. Organizing and structuring discourses for assigning painful trade-offs goes beyond the good intention to have all relevant stakeholders

involved in decision making. The mere desire to initiate a two-way-communication process and the willingness to listen to stakeholder concerns are not sufficient. Discursive processes need a structure that assures the integration of technical expertise, regulatory requirements, and public values. These different inputs should be combined in such a fashion that they contribute to the deliberation process the type of expertise and knowledge that can claim legitimacy within a rational decision making procedure (von Schomberg 1995). It does not make sense to replace technical expertise with vague public perceptions, nor is it justified to have the experts insert their own value judgments into what ought to be a democratic process.

Decision analytic tools can be of great value for structuring and assigning complex trade-offs. They can provide assistance in problem structuring, in dealing with complex scientific issues and uncertainty, and in helping a diverse group to understand disagreements and ambiguity with respect to values and preferences. Decision analysis tools should be used, however, with care. They do not provide an algorithm to reach an answer as to what is the best decision. Rather, decision analysis is a formal framework that can be used for trade-off analysis and risk handling to explore difficult issues, to focus debate and further analysis on the factors most important to the decision, and to provide for increased transparency and more effective exchange of information and opinions among the process participants. The basic concepts are relatively simple and can be implemented with a minimum of mathematics (Hammond et al. 1999).

Many risk management agencies are already making use of decision analysis tools. This, however, need further refinement. It is necessary to use these tools in the context of an iterative, deliberative process with broad participation by the interested and affected parties to the decision. The analytical methods, the data and judgment, and the assumptions, as well as the analytical results should be readily available and understood by the participants. Both the risk management agencies and the interested groups within the public that government agencies interact with on risk management decisions should all gain experience with these methods.

With respect to hormesis it is ethically mandated that potential beneficial aspects of low exposure to potentially hazardous material are incorporated in the risk-benefit balancing procedure. The potential harm done by pollutants do not justify the invocation of a categorical principle. Minimisation of risk is not required if health benefits are also at stake. Society needs to find an informed consent on the threshold of risk below compensation of goods is legitimate and morally justified. Such a threshold can be defined context-specifically but any human action associated with potential health impacts makes such an acceptability judgment – implicitly or explicitly. Incorporating hormesis into risk management forces regulators to make such thresholds explicit. Once a risk is below this threshold all positive and negative impacts are subject to a relative balancing towards reaching a final judgment on acceptability and necessary risk management options. This balancing of risk cannot be reduced to body counts: equity issues, context specific circumstances (voluntary or involuntary exposure, for example), avoidability of risks, the nature of vulnerable groups and many other factors need to be taken into account. Such a complex weighing exercise is best performed by an analytic-deliberative process by which the best available knowledge of impacts (including their uncertain ties) is fed into a deliberating body of individuals representing all sides of the debate. Such a debate would be inspired by the consensual and procedural school of ethics in which rational discourse seen as the most suitable instrument to come to a morally superior conclusion when facing conflicting values and principles. If such discourses were made effective in regulatory decision making, the debate about hormesis could act as a catalyst for needed regulatory reform.

7. BIBLIOGRAPHY

Akademie der Wissenschaften: *Umweltstandards.* (de Gruyter: Berlin 1992)

Albert, H.: *Traktat über kritische Vernunft.* (Mohr: Tübingen: 1991)

Asselt, van M.B.A.: *Perspectives on Uncertainty and Risk* (Kluwer: Dordrecht and Boston 2000).

Beck, U.: "Weltrisikogesellschaft, Weltöffentlichkeit und globale Subpolitik. Ökologische Fragen im Bezugsrahmen fabrizierter Unsicherheiten". In: A. Diekmann and C. Jaeger, (Ed.): *Umweltsoziologie. Kölner Zeitschrift für Soziologie and Sozialpsychologie* Sonderheft 36 (Westdeutscher Verlag: Opladen 1996), 119–147.

Calabrese, E.J. and Baldwin, L.A.: "Radiation Hormesis: The Demise of a Legitimate Hypothesis" , *Hum. Exp. Toxicology,* 19, (2000), 76-84

Calabrese, E.J. and Baldwin, L.A.: "The Frequency of U-Shaped Dose Responses in the Toxicological Literature", *Toxicological Sciences,* 61, (2001), 330-338

Calabrese, E.J.; Baldwin, L.A. and Holand, C.D.: "Hormesis: A highly Generalizable and Reproducible Phenomenon with Important Implications for Risk Assessment", *Risk Analysis,* 19, (1999), 261-281

Dodgson, J.; Spackman, M.; Pearman, A.; Phillips, L. D.: *Multi-Criteria Analysis: A Manual*. Department of the Environment, Transport and the Regions (London School of Economics: London 2000).

Dybing, E.; Doe, J.; Groten, J.; Kleiner, J.; O'Brien, J.; Renwick, A.G.; Schlatter, J.: Steinberg, P.; Tritscher, A.; Walker, R.; Younes, M.; Hazard Characterization of Chemicals in Food and Diet: Dose Response, Mechanisms and Extrapolation Issues. *Food Chemical Toxicology,* 40 (2003), 237-282

EFSA (European Food Safety Authority): Opinion of the Scientific Committee on a Request from EFSA Related to a Harmonised Approach for Risk Assessment of Substances Which are Both Genotoxic and Carcinogenic, Request No EFSA-Q-2004-020, *EFSA Journal,* 282 (2005), 1-31

Ehrenfeld, D.: *Beginning Again: People and Nature in the New Millenium*. (Oxford University Press: New York 1993).

Einhorn, H.J. and Hogarth, R.M.: "Confidence in Judgment: Persistence of the Illusion of Validity." *Psychological Review,* 85 (1978), 395-416.

Fischhoff, B.: "Hindsight Versus Foresight: The Effect of Outcome Knowledge on Judgment under Uncertainty." *Journal of Experimental Psychology: Human Perception and Performance,* 1 (1975), 288-299.

Frankena, W. K.: *Ethics.* (Prentice Hall: Englewood Cliffs 1963).

Galert, T.: *Biodiversität als Problem der Naturethik.* Literaturreview und Bibliographie. Graue Reihe Nr. 12. Bad Neuenahr-Ahrweiler: Europäische Akademie zur Erforschung von Folgen wissenschaftlich-technischer Entwicklungen. (European Academy: Neuenahr 1998)

Gethmann, C. F. (1991): "Ethische Aspekte des Handelns unter Risiko". In: M. Lutz-Bachmann (Ed.): *Freiheit und Verantwortung.* (de Gruyter: Berlin 1991), 152–169.

Gethmann, C. F.: "Rationale Technikfolgenbeurteilung". In: Grundwald, A. (Ed.): *Rationale Technikfolgenbeurteilung. Konzepte und methodische Grundlagen*. (Springer: Berlin, Heidelberg, New York: 1998), pp. 1–7.

Goodwin, P. and Wright, G.: *Decision Analysis for Management Judgement.* (Wiley: London 2004)

Gowdy, J.: "The Value of Biodiversity". *Land Economics* 73, 1 (1997), 25–41.

Habermas, J.: Vorbereitende Bemerkungen zu einer Theorie der kommunikativen Kompetenz. In: Habermas, J. and Luhmann, N. (Ed.): *Theorie der Gesellschaft oder Sozialtechnologie. Was leistet die Systemforschung?* (Suhrkamp: Frankfurt/M. 1972), 101–141.

Habermas, J.: *Theorie des kommunikativen Handelns*. (Suhrkamp: Frankfurt/M. 1981).

Habermas, J.: *Moralbewußtsein und kommunikatives Handeln*. (Suhrkamp: Frankfurt/M. 1991).

Hammond, J.; Keeney, R. and Raiffa, H.: *Smart Choices: A Practical Guide to Making Better Decisions* (Harvard Business School Press: Cambridge 1999).

Hartwich, H.-H. and Wewer, G. (Ed.): *Regieren in der Bundesrepublik*. Band 3: Systemsteuerung und "Staatskunst": Theoretische Konzepte und empirische Befunde. (Leske und Budrich: Opladen 1991).

Höffe, O.: *Politische Gerechtigkeit. Grundlegung einer kritischen Philosophie von Recht und Staat.* (Suhrkamp: Frankfurt/M. 1987)..

Höffe, O.: *Immanuel Kant.* (Beck: München 1992).

Honnefelder, L.: "Welche Natur sollen wir schützen?" *GAIA,* 2,5 (1993), 253–264.

Hubig, C.: *Technik- und Wissenschaftsethik. Ein Leitfaden.* (Springer: Berlin, Heidelberg, New York:1993)

IRGC (International Risk Governance Council): *Risk Governance: Towards an Integrative Approach.* White Paper No. 1, written by Ortwin Renn with an Annex by Peter Graham (IRGC: Geneva 2005)

Jonas, H.: *Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation*. (Suhrkamp: Frankfurt/M. 1979),.

Jasanoff, S.: "Beyond Epistemology: Relativism and Engagement in the Politics of
Science." *Social Studies of Science,* 26, 2 (1996), 393-418.

Kahneman, D. and Tversky, A., "Judgment under Uncertainty: Heuristics and Biases." *Science,* 185, (1974), 1124-1131.

Klinke; A. and Renn, O.: "A New Approach to Risk Evaluation and Management: Risk-Based, Precaution-Based and Discourse-Based Management." *Risk Analysis,* 22, 6 (December 2002), 1071-1994.

Lee, T.R.: "The Public Perception of Risk and the Question of Irrationality". In: Royal Society of Great Britain (Ed.): *Risk Perception* , Vol. 376 (Royal Society: London 1981), 5-16.

Lindbloom, C.: "The Science of Muddling Through." *Public Administration Review,*19, (1959), 79–99.

Lindbloom, C.: *The Intelligence of Democracy. Decision Making through Mutual Adjustment.* (Free Press: New York: 1965)

Luhmann, N.: *Legitimation durch Verfahren*. (Suhrkamp: Frankfurt/M. 1983),.

Merkhofer, M.W.: "Comparative Analysis of Formal Decision-Making Approaches". In: V.T. Covello, J. Menkes. und J. Mumpower (Eds.): *Risk Evaluation and Management* (Plenum: New York 1984), 183-220.

Mittelstraß, J.: "Gibt es eine Letztbegründung?" In: Janich, P. (Ed.): *Methodische Philosophie.* (Econ: Düsseldorf 1984), 12–35.

Mittelstraß, J: *Leonardo-Welt.* (Suhrkamp: Frankfurt/M. 1992),.

National Research Council, Committee on Institutional Means for Assessing Risk to Public Health, *Risk Assessment in the Federal Government: Managing the Process* (National Academy Press: Washington, D.C. 1983). http://books.nap.edu/catalog/366.html

National Research Council, Committee on Risk Characterization, *Understanding Risk: Informing Decisions in a Democratic Society.* (National Academy Press: Washington, D.C.1996a). http://books.nap.edu/catalog/5138.html

Nida-Rümelin, J.: "Ethik des Risikos". In: J. Nida-Rümelin (Ed.): *Angewandte Ethik. Die Bereichsethiken und ihre theoretische Fundierung.* (Kröner: Stuttgart 1996a), 806–831.

Nida-Rümelin, J.: "Theoretische und Angewandte Ethik: Paradigmen, Begründungen, Bereiche". In: J. Nida-Rümelin (Ed.): *Angewandte Ethik. Die Bereichsethiken und ihre theoretische Fundierung.* (Kröner: Stuttgart 1996b), 2–85.

Ott, K. : *Ökologie und Ethik. Ein Versuch praktischer Philosophie.* Ethik in den Wissenschaften. Band 4. (Attempto: Tübingen 1993).

Ott, K.: "Wie ist eine diskursethische Begründung von ökologischen Rechts- und Moralnormen möglich?" In: K. Ott (Ed.): *Vom Begründen zum Handeln. Aufsätze zur angewandten Ethik* (Attempto: Tübingen 1996), 86–128.

Randall, A.: "What Mainstream Economists Have to Say about the Value of Biodiversity." In: E.O.Wilson (Ed.): *Biodiversity.* (Academy Press: Washington, D.C.:1988), 217–223.

Randall, A. and Farmer, M. C. (1995): "Benefits, Costs, and the Safe Minimum Standard of Conservation". In: Bromley, D. W. (Ed.): *The Handbook of Environmental Economics. (*Blackwell: New York, Oxford 1995), 26–44.

Renn, O.: "Die Wertbaumanalyse. Ein diskursives Verfahren zur Bildung und Begründung kollektiv verbindlicher Bewertungskriterien". In: A. Holderegger (Ed.): *Ökologische Ethik als Orientierungswissenschaft.* (Universitätsverlag: Freiburg 1997), 34–67.

Renn, O.: "The Challenge of Integrating Deliberation and Expertise: Participation and Discourse in Risk Management." In. T. McDaniels and M.J. Small (Eds); *Risk Analysis and Society. An Interdisciplinary Characterization of the Field.* (Cambridge University Press: Cambridge 2004), 289-366.

Revermann, C.: "Was heißt hier Bioethik?" *TAB-Brief* 15 (1998), 9–16.

Rock, M.: "Theologie der Natur und ihre anthropologisch-ethischen Konsequenzen". In: D. Birnbacher (Ed.): *Ökologie und Ethik.* (Reclam: Stuttgart 1980).

Ropohl, G.: "Ob man die Ambivalenzen des technischen Fortschritts mit einer neuen Ethik meistern kann?" In: H. Lenk and M. Maring (Eds.): *Technikverantwortung.* (Campus: Frankfurt/M., New York: 1991), 48–78.

Schmitz, P.: *Ist die Schöpfung noch zu retten? Umweltkrise und christliche Verantwortung.* (Echter: Würzburg 1985).

Shrader-Frechette, K.: *Environmental Ethics*. (Boxwood Press: Pacific Grove 1988)

Shrader-Frechette, K.S.: *Risk and Rationality. Philosophical Foundations for Populist Reforms* (University of California Press: Berkeley 1991).

Slovic, P.; Fischhoff, B. and Lichtenstein. S.: "Behavioral Decision Theory." *Annual Review of Psychology,* 28 (1977), 13-28.

Stebbing, A.R.D.: Hormesis- the Stimuzlation of Growth by Low Levels of Inhibitors, Sci.Total Environment, 22 (1982), 213-234

Stebbing, A.R.D.: "A Theory for Growth Hormesis". *Mutat. Res.,* 403, (1998), 249-258

Stern, P. C. and Fineberg, V.: *Understanding Risk: Informing Decisions in a Democratic Society.* (National Academy Press: Washington 1996).

Szejnwald-Brown, H. S., Derr, P., Renn, O. and White, A. L.: *Corporate Environmentalism in a Global Economy. Societal Values in International Technology Transfer.* (Quorum Books: Westport 1993).

Tversky, A. and Kahneman, D.: "The Framing of Decisions and the Psychology of Choice." *Science,* 211 (1981), 453-458.

US EPA (US Environmental Protection Agency): *The Use of Benchmark Dose Approach in Health Risk Assessment.* EPA/630/R-94/007. Risk Assessment Forum (EPA: Washington, D.C. 1995).

Van der Sluijs, J.P., Risbey, J.S., Kloprogge, P., Ravetz, J.R., Funtowicz, S.O., Corral Quintana, S., Guimaraes Pereira, A., De Marchi, B., Petersen, A.C., Janssen, P.H.M., Hoppe, R. and Huijs, S.W.F.: *RIVM/MNP Guidance for Uncertainty Assessment and Communication. Report No. NWS-E-2003-163* (Copernicus Institute for Sustainable Development and Innovation and Netherlands Environmental Assessment Agency: Utrecht and Bilthoven 2003)

Viscusi, W.K.: "Risk-Risk Analysis," *Journal of Risk and Uncertainty,* 8 (1994), 5-18.

von Schomberg, R.: "The Erosion of the Valuespheres. The Ways in Which Society Copes with Scientific, Moral and Ethical Uncertainty." In: R. von Schomberg (Ed.): *Contested Technology. Ethics, Risk and Public Debate* (International Centre for Human and Public Affairs: Tilburg 1995), 13-28.

Vollmer, H.: "Akzeptanzbeschaffung: Verfahren und Verhandlungen". *Zeitschrift für Soziologie,* 25, 2 (1996), 147–164.

WBGU – Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen: *Welt im Wandel: Erhaltung und nachhaltige Nutzung der Biosphäre.* Jahresgutachten 1999. (Springer: Berlin, Heidelberg, New York:2000)

WBGU -Wissenschaftlicher Beirat der Bundesregierung Globale Umweltveränderungen: *Sondergutachten: Die Welt im Wandel: Umwelt und Ethik.* (Metropolis: Marburg 1999)

Weinrich, H.: "System, Diskurs, Didaktik und die Diktatur des Sitzfleisches". *Merkur,* 8, (1972), 801–812.

Wellmer, A.: "Konsens als Telos der sprachlichen Kommunikation?" In: H.-J. Giegel (Ed.): *Kommunikation und Konsens in modernen Gesellschaften.* (Suhrkamp: Frankfurt/M: 1992), 18–30.

Wiener, J.B.: "Managing the Iatrogenic Risks of Risk Management," *Risk: Health Safety & Environment,* Vol. 9 (1998), 39-83.

Wilson, E. O.: *Biophilia: The Human Bond with Other Species.* (Harvard University Press: Cambridge/MA; 1984).

Wolters, G.: "'Rio' oder die moralische Verpflichtung zum Erhalt der natürlichen Vielfalt. Zur Kritik einer UN-Ethik". *GAIA* 4, 4 (1995), 244–249.

Zilleßen, H.: "Die Modernisierung der Demokratie im Zeichen der Umweltproblematik". In: H. Zilleßen: P.C. Dienel, and W. Strubelt. (Eds.): *Die Modernisierung der Demokratie.* (Westdeutscher Verlag: Opladen 1993), 17–39.

Zimmerli, W. C.: "Wandelt sich die Verantwortung mit technischem Wandel?" In: H. Lenk, and G. Ropohl (Eds.): *Technik und Ethik.* (Reclam: Stuttgart 1993), 92–111.

[1] It should be noted that this chapter draws from material that the author has compiled for the German Scientific Council for Global Environmental Change and that has been partially published in German in a special report of the Council (WBGU 2001). The last section on decision making has borrowed material from an unpublished background document on decision making and risk management that Dr. Warner North and the author had prepared for the US National Academy of Sciences.

THE ETHICS OF HORMESIS – NO FUSS?

Per Sandin
Dept. of Philosophy and the History of Technology
Royal Institute of Technology – KTH
Teknikringen 78B
SE-100 44 Stockholm, Sweden
sandin@infra.kth.se

INTRODUCTION

There are different positions in ethics, and the debate on the ethics of risk is no exception. There are two main strands in this debate: Consequentialist approaches and non-consequentialist approaches.

According to a consequentialist approach to ethics, an action's moral status is determined by the consequences of the action. Certain areas of standard risk management have a consequentialist basis: It is the consequences that count. More precisely, they amount to applied forward-looking utilitarianism, with utility understood as the avoidance of harms to human health. (By 'forward-looking' I mean that it is used for choosing between possible courses of action, rather than evaluate them afterwards, for instance for purposes of assigning responsibility or blame.) Utilitarianism tells us to maximise *overall* benefit. In classical utilitarianism, this is captured in the slogan of 'greatest happiness for the greatest number'. For risk management, the slogan could perhaps be rephrased: 'Least harm to the greatest number'. Obviously, even for very simple, everyday actions it is difficult or impossible to know beforehand what consequences those actions will have. Thus a forward-looking utilitarian (and a risk manager) would typically recommend choosing an action that maximises *expected* utility. Or, as in the case of rule consequentialism, recommend a rule that, if followed, will lead to the best overall consequences in the long run.

On a non-consequentialist approach to ethics, on the other hand, an action's moral status is not completely determined by its consequences, even if consequences certainly can be one factor to be taken into account in the evaluation of an action. One example of a non-consequentialist approach to ethics is a rights-based approach, according to which actions that infringe individuals' rights are prohibited. Another example of a non-consequentialist approach is the Kantian approach, according to which an action's moral status is determined by whether it is performed on the basis of a maxim that is rationally universalizable.

Both approaches have their advantages and drawbacks. One of the cornerstones of utilitarianism, which standard risk management has inherited, is that utility is interpersonally comparable and aggregable, that is, one person's utility counts just as much as another's, and the overall sum of utility is what should be maximised. This means that an action that exposes a large number of people to a small risk is preferable to one that exposes a small number of people to a very large risk. For instance, suppose that we are deciding about whether we should (a) expose 200 people to 2 units of risk each, or (b) expose 10 people to 39 units of risk each. The total risk is smaller by ten units in case (b) than in case (a), and therefore, we should opt for (b). This conclusion is counterintuitive to many, since it conflicts with common ideas about fairness.

On the other hand, non-consequentialist approaches can also lead to counterintuitive conclusions. One such problem is clearly related to risk. Take, for instance, a rights-based approach that prohibits the infringement of the rights of individual persons. A plausible example of such a right would be the right to life. However, given that we are in a situation of risk, we cannot know whether a certain action will be an infringement of someone's right to life. On the other hand, we *can* know that many people will be exposed to a *risk* of having their right to life infringed. If exposing other people to such risks would be prohibited, with reference to individuals' rights, no matter how small the risk is, it would be very difficult to arrange a society.

Problems such as these have led to refinement of consequentialist and non-consequentialist approaches. In risk management and policy contexts, one often finds combinations of utilitarian ideas (maximising the overall good) and concerns for distributional issues (such as fairness) and protection for individuals. Thus, a strategy of 'maximising within constraints' seems intuitively appealing, and such systems have been in use.[i]

HORMESIS AND POLICY PROBLEMS

By 'hormesis' I will mean 'a dose-response phenomenon characterized by either a U-shaped or inverted U-shaped dose-response depending on the end-point measured.'[ii] An example of this phenomenon would be a substance, which in small doses protects those exposed from cardiovascular disease, but in high doses promotes cardiovascular disease. Alcohol might be such a substance. In the remainder of this article, I will consider three problems that the incorporation of considerations of hormetic effects into policy will give rise to: (1) The comparative smallness of hormetic effects; (2) the fine-tuning problem; and (3) the problem of aggregated action. I will argue that these three problems should lead us not to overemphasise the importance of hormesis for policy, and that they, if anything, points towards a non-consequentialist approach to the ethics of risk.

THE SMALLNESS OF HORMETIC EFFECTS

The first problem when pondering the ethical implications of hormesis is that hormetic effects seem comparatively small. Calabrese writes 'Regardless of plant, microbe, fish, rodent, in vitro or in vivo systems, the modest response remains perhaps the most significant feature of the hormetic dose-response phenomenon'. And 'the maximum stimulatory response is … often difficult to confidently distinguish from normal variation.' [iii]

At the same time, we know that there are large health effects from higher doses of several agents. As an example, consider once again alcohol. It may be that there are small beneficial effects from drinking, say, one glass of red wine per day. But we know that drinking a couple of bottles of wine per day has very bad effects on health. (Of course, apart from medical effects such as liver cirrhosis, diabetes, hypertonia etc, there are social effects of alcohol abuse as well: domestic violence, crime, traffic accidents due to drunk driving, and so on.)

In the case of alcohol, overconsumption is a serious societal problem. Thus, in setting priorities, it seems that possible beneficial low-dose effects of alcohol is not a key issue.

Two things may be noted, though: First, there may certainly be agents where the discrepancy between beneficial low-dose effects and detrimental high-dose effects is less conspicuous than in the case of alcohol. Suppose, for instance, that a chemical substance has beneficial low dose effects and detrimental high-dose effects, but that the high-dose effects only occur at extremely high levels, to which no one is ever exposed. In that case, hormetic effects might be more relevant. Second, there is a distributional issue. Most people who drink alcohol drink moderately, while comparatively few are heavy drinkers. (According to a report issued by the Swedish National Institute of Public Health, 8.3 % of the male population of Sweden between 16 and 75 years of age could be classified as heavy drinkers, defined as consuming more than 30 grams of alcohol per day.)[iv] If very few people are exposed to high doses, with ensuing large detrimental effects, and a very large number of people are exposed to small doses with small beneficial effects, then, arguably, the beneficial low-dose effects should be given serious consideration. The sum of the beneficial low-dose effects may, at least on a maximising consequentialist approach, outweigh the harm done to a few. However, this is a position that certainly will be highly controversial.

THE FINE-TUNING PROBLEM

The second problem is what may be termed the *fine-tuning problem*. Consider a highly simplified example, again involving alcohol. Suppose that alcohol displays a hormetic dose-response curve, and that the following holds: One glass of wine per day is good for your health, so that one glass per day is slightly better than total abstention. Two glasses of wine per day is equal in health effects to total abstention. At three or more glasses of wine per day, negative health effects take over, so that drinking more is increasingly detrimental to health.

In order to bring about the best consequences, in this case in the form of health effects, we should design a policy that is conducive to people's consuming one drink per day, neither more nor less. (Other things being equal.) Alcohol consumption should thus be *fine tuned*. The problem is that fine tuning is significantly more difficult to achieve than, say, plain reduction of exposure. Compare the requirement of fine-tuning with the ALARA principle, which states that exposure should be kept 'as low as reasonably achievable'. This principle has its origin in radiation protection, but has been discussed in other areas as well. The benefit of a principle such as ALARA is its comparative simplicity. The same holds for absolute dose limits, which say that a certain level of exposure must not be exceeded.

In summary: If hormetic effects are to be given a place in policy, the fine-tuning problem would call for a more complex policy. This would probably lead to policy recommendations that, in being more complex, also are less transparent. Simplicity and transparency are certainly not the only criteria that should be used in evaluating policy measures, but they are very important for whether the measures will be considered legitimate.

AGGREGATED ACTIONS

The third problem is that of aggregated actions. This is a well-known problem in consequentialist ethics. The consequences of my actions are determined not only but what I do, but also by what others do, and what I do at other times.[v] An action which in itself would have beneficial (or non negligibly detrimental) consequences, can be part of a *set* of actions, which *together* have very negative consequences.

Take the alcohol example again. The endpoint 'health', assuming that it can be given a reasonably precise definition, is obviously affected not only by alcohol, but by a whole range of different agents as well: food, cigarettes, background radiation, environmental pollutants, and so on. The beneficial effects of low doses of alcohol—supposing that there are such effects—can be cancelled out as well as reinforced by exposure to other agents. For instance, suppose that it can be established that agents A_1, A_2 and A_3 each have beneficial low dose effects. However, the question whether exposure to low doses of all three agents simultaneously is beneficial is not answered thereby, but needs independent investigation. This problem, which has not gone unnoticed,[vi] further complicates the question of how policy recommendations would be affected if the hormesis concept can be shown to be ubiquitous.

A similar problem occurs in the context of *de minimis* risk.[vii] *De minimis* risks are risks that are so small that they are negliglible, and that they thus can be disregarded in the decision process.[viii] Suppose that, as has been actually suggested, a lifetime risk of one in a million is *de minimis* and thus can be disregarded. However, there may be several such risks, which taken together make up a significant level of total risk.[ix] There is no obvious solution to this problem, and it is highly relevant also in the hormesis case.

CONCLUSION

To conclude, then: What should we make of hormesis, from an ethical point of view? A tentative answer is: not too much. The three problems discussed above indicates this: First, hormetic effects are small and it is therefore doubtful whether they should be given much weight, since there are more pressing problems. Second, benefiting from them requires fine-tuning, which calls for more complex and possibly less transparent policy measures. Third, attention needs to be

given to the effects of sets of actions, such as actions resulting in simultaneous exposure to a number of agents.

These three problems—the smallness of effects, the fine-tuning problem and the problem of aggregated actions—point in the direction of not giving too much weight to hormesis in policy contexts.

If some weight is given to them, though, the problems point in another direction as well: towards a non-consequentialist approach to risks. The reason is the following: Problems posed by hormesis are primarily about the difficulty of assessing consequences of exposure to different types of agents. If our previous hypotheses do not hold – for instance, if the linear non-threshold hypothesis about radiation exposure and cancer incidence is false and low-dose exposure to radiation actually prevents cancer – then our usual models for assessing consequences are no longer valid, and that complicates things, if consequences are what counts, morally. A non-consequentialist approach to ethics, on the other hand, would not require revision to the same extent. Therefore, if there is some importance in the hormesis concept, it would, arguably, provide *one* reason for favouring a non-consequentialist over a consequentialist approach to the ethics of risk. This does obviously not suffice to establish the superiority of non-consequentialist approaches. That discussion, however, has to be postponed for the time being.

REFERENCES

[i] Wikman-Svahn P, Peterson M, Hansson SO. Principles of protection: a formal approach for evaluating dose distributions. *Journal of Radiological Protection* 2006; **26**: 69-84.

[ii] Calabrese EJ, Baldwin L. Hormesis: U-shaped dose responses and their centrality in toxicology. *TRENDS in Pharmacological Sciences* 2001; **22**: 285-291.

[iii] Calabrese EJ. Paradigm lost. Paradigm found: The re-emergence of hormesis as a fundamental dose-response model in the toxicological sciences. *Environmental Pollution* 2005; **138**: 379-412.

[iv] Leifman H. Alkohol i Sverige – konsumtion och dryckesmönster. In Andréasson S, Allebeck P (eds.) *Alkohol och hälsa: En kunskapsöversikt om alkoholens positiva och negativa effekter på vår hälsa*. Report R2005:11, Swedish National Institute of Public Health, Stockholm: 15-35. [In Swedish]

[v] Parfit D. *Reasons and Persons*. Oxford University Press, 1987.

[vi] Thayer KA, Melnick R, Burns K, Davis D, Huff J. Fundamental flaws of hormesis for public health decisions. *Environmental Health Perspectives* 2005; **113**, 1271-1276.

[vii] Shrader-Frechette K. Technological risk and small probabilities. *Journal of Business Ethics* 1985; **4**: 431-446.

[viii] Sandin P. Naturalness and de minimis risk. *Environmental Ethics* 2005; **27**: 191-200.

[ix] Wikman P. Trivial risks and the new radiation protection system. *Journal of Radiological Protection* 2004; **24**, 3-11.

IDEOLOGICAL TOXICOLOGY:

INVALID LOGIC, SCIENCE, ETHICS ABOUT LOW-DOSE POLLUTION

Kristin Shrader-Frechette
O'Neill Family Professor
Department of Biological Sciences and Department of Philosophy
University of Notre Dame
100 Malloy Hall
Notre Dame, IN 46556
kshrader@nd.edu
574-631-2647

IDEOLOGICAL TOXICOLOGY: INVALID LOGIC, SCIENCE, AND ETHICS OF LOW-DOSE POLLUTION

If scientists rely on assumptions rather than logic, empirical confirmation, and falsification, they are no longer doing science, but ideology – which is, by definition, unethical. As a recent US National Academy of Sciences report put it, "bad science is always unethical."[1] This paper discusses several ways in which toxicologists can fall into ideology – bad, therefore unethical, science.

In part because of the increasing expense of pollution control, some toxicologists have been re-examining pollution dose-response curves that are non-monotonic, i.e., curves in which the direction of some response changes with increasing or decreasing dose.[2] Ethanol is a classic example of a non-monotonic dose-response curve, because moderate drinking is associated with lower risks of heart disease, while heavy drinking is associated with higher risks.[3,4] If some low-dose pollutants exhibit adaptive or "beneficial effects,"[5] this might suggest re-thinking pollution regulations which presuppose linear, no-threshold (LNT) dose-response curves.

OVERVIEW

As illustrated by the case of ethanol, claim H is that, for some biological endpoints, low-dose toxins and carcinogens exhibit hormesis, a "beneficial,"[5] or "adaptive, response characterized by biphasic dose responses" and resulting from "compensatory biological processes following an initial disruption in homeostasis."[6] From this uncontroversial claim H, however, the paper argues that some toxicologists invalidly infer HG (that H is "generalizable across biological model, endpoint measured, and chemical class"[7]) and HD (that "a strong case can be made for the use of hormesis [H] as a default assumption in the risk-assessment process"[2]). Evaluating HG and HD, this paper argues for 5 claims. While (1) H is true, (2) HG falls victim to several logical fallacies and therefore is logically, scientifically, and ethically invalid. (3) Because it relies on logical fallacies, confuses necessary and sufficient conditions, and violates at least 5 sets of ethical norms, HD is logically, scientifically, and ethically invalid. (4) Five remedies could help address HG-HD flaws and failure to adequately assess low-dose exposures. (5) Three objections to these criticisms of HG and HD are easily answered.

H IS SCIENTIFICALLY UNCONTROVERSIAL BECAUSE OF ITS LIMITED SCOPE

As many examples attest, claim H (that, for some biological endpoints, some low-dose toxins and carcinogens exhibit hormesis – a "beneficial,"[5] or "adaptive, response characterized by biphasic dose responses" and resulting from "compensatory biological processes following an initial disruption in homeostasis"[6]) is both true and uncontroversial. H is true and uncontroversial, however, largely because it requires so little: at least one non-monotonic effect, on one endpoint, from one pollutant, for one period of time. H would be satisfied if a pollutant caused cancer (one endpoint), but increased fingernail growth (another endpoint). Thus, low-dose cadmium satisfies H in reducing some tumors in some species and increasing growth in some plants, although tests on 8700 adults showed that low-dose cadmium is associated with excess prediabetes and diabetes, and animal tests showed pancreas damage, glucose dysregulation, and kidney damage.[8] Likewise, moderate drinking of 1.2 to 2.2 drinks per day satisfies H because it reduces mortality, yet it increases breast-cancer risk.[8] The upshot? Given the minimalist definition of H, when low-dose responses are beneficial for some endpoints but harmful for others, the response nevertheless satisfies H.

Indeed, the Calabrese-Baldwin conditions for H are so minimal that they call responses "hormetic,"[2] when (a) alleged H responses do not satisfy criteria for statistically significant changes from control. Thus, a not-statistically-significant change in incidence from 2 to 3, in a sample of 20, is called a 33-percent change, evidence of hormesis. Likewise, Calabrese-Baldwin use a study no-observed-adverse-effect level or NOAEL to assess H. Because sample size, statistical power, data variability, endpoint measured, duration of exposure, route of exposure, rate of exposure, and so on, affect study NOAEL, therefore (b) alleged H responses can be merely artifacts of factors like small sample size or data variability.[3,9]

Because scientific criteria for H are minimalist, not scientifically rigorous, instances of alleged H responses are easy to find. Yet they reveal almost nothing about total responses, net beneficial effects, lifetime responses, or all-endpoint effects – factors that are crucial to reli-

ably assessing the policy-relevance of alleged low-dose responses to toxins like TCDD (dioxin). Consider four methodological flaws in a two-year, low-dose TCDD test on rats, a test alleged to illustrate H, decreased tumor incidence.[10] *First*, the study covered only about two-thirds of the rats' lifespan, not the most vulnerable periods. If roughly 80 percent of human cancers are diagnosed in the last one-third of life,[11] and if the rat analogy holds for human lifespan/cancers, the study may have captured only 20 percent of cancers induced by TCDD, not total cancers. *Second*, although liver, lung, tongue, and nasal tumors increased in this study, while pituitary, uterine, mammary, pancreas, and adrenal tumors decreased, the study invalidly aggregated all tumors. Because no individual tumor response was non-monotonic, the alleged H response seems an artifact of invalid aggregation. *Third,* the study also ignored early mortality and confounders like lower body weights when it calculated tumor rates, relative to controls. *Fourth*, there may be a replication problem, because other TCDD studies (in primates) have shown a variety of low-dose adverse effects.[12] Despite these four methodological problems, the study has been used to allege H.[8,3]

GENERALIZING TO HORMESIS CLAIM HG: LOGICALLY, SCIENTIFICALLY, AND ETHICALLY INVALID

Given the lack of rigorous scientific conditions for (and thus the relative ease of) claiming an instance of H, there are obvious scientific problems with generalizations based on H. HG is the claim that H is "generalizable across biological model, endpoint measured, and chemical class,"[7] that "the hormetic model is not an exception to the rule [of linear, no-threshold or LNT dose-responses] – it is the rule"[13]

One indicator of HG's potential problems is that the classic cases from which HG is most often inferred, those of Calabrese and Baldwin,[2] include no epidemiological or field studies.[3] Yet these types of studies are precisely those in which conditions best mimic real-world exposure, and in which HG is most likely to be refuted.

Limited scientific information is another indicator of HG's problems. As a consequence, inferring HG, that claim H is true, "generalizable across…endpoint measured," often commits the fallacy of appeal to ignorance. This fallacy occurs when people assume that because no evidence refutes a claim, therefore it is true. They invalidly assume that the absence of some evidence (e.g., against HG) constitutes evidence of the absence (e.g., of data against HG). For instance, US National Academy of Sciences' studies have warned that, despite known higher sensitivities of children to pesticides and herbicides, and despite current regulations' not adequately protecting them, nevertheless data are inadequate to precisely define these higher sensitivities for children's neuro-developmental effects or endpoints.[14] Yet to posit HG, one must commit the fallacy of appeal to ignorance and assume that, despite scientific ignorance (e.g., about precise pesticide-herbicide effects on children's neuro-developmental endpoints), HG holds for all endpoints. Yet to confirm that HG holds, as adaptive across all endpoints, there must be evidence from large-sample, long-term, in-depth, all-endpoint studies. In the absence of such sophisticated studies – clearly not those typically used to assert H – HG proponents commit the fallacy of appeal to ignorance.

HG proponents also exhibit the inductive fallacy (also called the fallacy of invalid extrapolation or the fallacy of hasty generalization) when they generalize or extrapolate to all endpoints, groups, and time-periods, on the basis of only a few endpoints, population-subgroups, or time-periods. The earlier cases of cadmium and ethanol illustrate why the HG extrapolation (to all endpoints) constitutes an inductive fallacy. HG extrapolation to all individuals and population-subgroups likewise is problematic because of genetic and lifestyle differences, e.g., certain medications can affect responses to toxins. HG extrapolation to all age groups is particularly questionable because of children's vulnerability. Some pharmaceuticals have half-lives that are 3 to 9 times longer in neonates than in adults, and neonates may have elimination half-lives that are more than 10 times longer than adults. In the case of alcohol, for example, while maternal drinking of 1.2 to 2.2 drinks per day may have beneficial effects on the mother, only 0.5 drinks per day have been associated with adverse behavioral and developmental effects on the fetus. Even apart from adult-child differences, among adults responses to pesticides, for example, may vary significantly because of factors like seven-fold differences in levels of detoxifying enzymes.[8]

Ignoring the endpoint/individual/age and other differences just illustrated, HG proponents' inductive fallacies are especially objectionable because they explicitly and harshly criticize those who extrapolate from high-dose to low-dose responses. Consistency therefore requires HG proponents to practice what they preach. They must avoid invalid extrapolations from some biological endpoints to all endpoints; from adult, pure-bred, homogenous animal populations of toxicological studies to non-adult, non-pure-bred, and heterogeneous members of human populations; and from some adaptive responses to net adaptive responses. They also must avoid extrapolating (purely on the basis of a simple, quantitative, low-dose measurement) to dose effects that are determined not only by quantity, but also by when the dose is received, who receives it, what is her health and nutritional status, how it is received (e.g., the dose rate), and with what it is received, e.g., other exposures. In using the inductive fallacy to extrapolate in all these ways, HG proponents not only "trim" the relevant dose data that are most likely to show HG false but also err in the same ways as those they criticize.

Apart from logic and scientific method, there are good biological reasons that individual, low-dose, adaptive responses are unlikely to be generalizable, overall, as adaptive – as HG requires. One reason is that, as Calabrese and Baldwin recognize,[2] hormesis effects are likely "overcompensations in response to disruptions in homeostasis." But when organisms overcompensate in order to respond to threats or disruptions, they pay a price. There is no free lunch. The adrenalin-rushes that are temporarily adaptive are, over the long term, maladaptive. Likewise, while overcompensatory responses to some toxin obviously have some adaptive benefits, they also obviously have metabolic costs – costs that, over the long term, may be harmful. HG proponents ignore these biological facts.

Because HG proponents fall victim to inductive fallacies and appeals to ignorance when they generalize to all endpoints, all responses, all subjects, all ages, and all exposure conditions, they beg the question of whether HG is true or not. Instead of offering detailed empirical evidence for all of these generalizing inferences, they merely assume it. Moreover, because HG is scientifically and logically invalid, it also is ethically invalid. A recent US National Academy of Sciences analysis made a similar point: "bad science is always unethical."[1] Discussing "studies in which people…make the case for setting a less stringent [pollutant] exposure standard," the academy authors warned that, because "studies that do not meet the highest scientific and ethical standards" have great potential to mislead scientists and regulators, they "should not be…accepted…as input to the regulatory decision-making process."[1]

USING HD IN REGULATION: LOGICALLY, SCIENTIFICALLY, AND ETHICALLY INVALID ARGUMENTS

Consider the consequences of preceding arguments for the claim HD, that "a strong case can be made for the use of hormesis [H] as a default assumption in the risk-assessment process."[2] Obviously, if the generalization HG is logically, scientifically, and ethically invalid, using it to infer HD is also invalid. Risk-assessment policy and regulation, like HD, should not be based on invalid, therefore unethical, science.

However, even if HG were true for most endpoints (and there is much evidence that it is not) – this would not justify HD – that is, it would not justify using HG as a default position in risk assessment and regulation. For one thing, even if hormetic, adaptive responses to a pollutant held across most endpoints, as HG posits, this fact constitutes only necessary, not sufficient, conditions for accepting HD. In addition, at least 5 other necessary conditions – ethical conditions – would have to be met, in order to accept HD.

One ethical condition is (a) that HD would have to represent an adequately health-protective stance, in the face of uncertainty about precise risks.[15,16,17] Because default rules like LNT and HD are used in situations of uncertainty, their acceptance is not a purely scientific decision. Rather, their acceptance is an ethical decision – about how much risk people will accept, who should take those risks, whether the benefits are worth it, and so on – given uncertainty about the possible ramifications of the risks. Hence, promoting an essentially ethical/policy claim, HD, largely on the basis of an allegedly scientific argument, HG, is invalid because HG-HD proponents attempt to deduce an ethical "ought" (HD) from a nonethical or allegedly scientific "is" (HG).[18] Yet, solely from what is the case, allegedly HG, it is never valid to deduce what ought to be the case, allegedly HD. To make this deduction is to commit the is-ought fallacy in ethics.

In addition to establishing HG scientifically and avoiding the is-ought fallacy in ethics, HD proponents would at least have to argue ethically (b) that it is equitable, compensable, just, and so on, to impose HD's possible risks on citizens; (c) that risk bearers should and would give informed consent to this HD default rule; (d) that the rule is operationalizable; and (e) that it satisfies basic rules of biomedical ethics.[17] No HD proponents have arguments meeting these 5 standard ethical conditions for risk imposition.

Moreover, several reasons suggest HD could not meet ethical condition (c), for consent. One reason is that people generally agree to bear uncertain risks, like those associated with a default rule, when they get something in return. Breast-cancer patients may take tamoxifen – despite its uncertain but excess risks of thrombosis, stroke, uterine hyperplasia, uterine cancer, and uterine sarcoma,[8] because they get something in return, reduced risk of breast-cancer recurrence. In fact, virtually all pharmaceuticals impose one risk, in exchange for reducing another risk. Ethics handles such mixed-risk pharmaceutical cases through informed consent.[17] Hence, even if people were adequately informed about HD risks, they likely would not consent, particularly if their children could be most at risk, or if they received nothing in return. As later paragraphs show, if the main HD beneficiaries are polluters, not the people who would bear most of the risks, HD is unlikely to satisfy the consent condition.

Likewise, HD proponents seem unable to meet ethical condition (d) because operationalizing and applying HD to the real world is impossible. Yet by the "ought implies can" rule, people can never be required to do what is impossible for them to do.[19] People cannot be required to spread their wings to fly, to rescue someone in the ocean, because it is impossible for people to spread their nonexistent wings. To say they "ought" to perform such a rescue implies they "can." If they cannot, logically they have no obligation to do so. Calabrese, Cook, and Baldwin forget this fundamental ethical rule – and its logical consequence.[1,13,20] Instead they repeatedly urge regulatory and risk-assessment changes, so as to take account of what is impossible for most people, viz., having total effects that are low-dose. But regulators and assessors need/ought not make such changes to HD, because they cannot. They cannot for two reasons, (i) because each person's exposure cannot be titrated, to achieve a total exposure that is narrow and low-dose, and (ii) because typical multiple doses of pollutants drive total exposures beyond low doses. To see these impossibilities, consider that Calabrese and Baldwin claim that maximal low-dose hormetic response occurs on average at a dose fivefold below the NOAEL.[21] If so, it logically follows that simultaneous exposure to 5 equally potent hormetic agents, each at one-fifth the NOAEL, could move the victim from the low-dose range to that of adverse effects. Yet it would be impossible, given a lifetime of fluctuating exposures and concentrations, to titrate each person's exposure to achieve a narrow, hormetic-exposure range.[8] Repeated US Environmental Protection Agency and Centers for Disease Control studies have shown that all US citizens have received doses of hundreds of chemicals whose residues are measurable in their blood or tissue.[8,3] Immunological evidence also shows that the combination of many low-dose effects is not always additive but synergistic, as when people are exposed to TCDD and numerous dioxin-like compounds, or to radon and smoking, to asbestos and smoking, to alcohol and smoking; more and more exposures add to the total immunologic and estrogenic burden.[22]

Likewise, although Calabrese and others repeatedly claim (HG) that low-dose radiation is adaptive, hormetic, or beneficial,[4,21,23,24] their claim contradicts all classic, consensus-position radiation studies, like those of the US National Academy of Sciences, which affirm LNT. (Only radiation studies whose authors have obvious conflicts of interest, like those of the French, reject LNT, but these conflicted studies are rejected by the global scientific community).[25,26] Yet even if HG were true for radiation, HD would not be operationalizable in the radiation case, any more than it is for the chemical case. Because all scientists agree that ionizing-radiation doses are cumulative, by the time a child is born, she has already received more than a low dose.[26] Thus, even if HG were true, because of the impossibility that most people's total doses of radiation or chemicals were low, and because of the impossibility of titrating such low-dose exposures, the ought-implies-can rule means that HD cannot meet (d) the operationalizability problem. HG is thus an irrelevant artifact, inapplicable to HD's real-world policymaking. That is, apart from its ethical problems, the inference from HG (about low doses) to HD (about total real-world responses, that are almost never low dose) commits the logical fallacy of irrelevant thesis and therefore is also unethical.

HD proponents likewise are unable to meet ethical condition (e), adherence to basic norms of biomedical ethics, as set out in classic statements like the Nuremburg Rules, the Belmont Report, the Helsinki Report, and the Common Rule of the US.[17] These all require that, before any risk is imposed on a subject, she must give free informed consent to that risk, part of which involves full risk disclosure and full risk understanding.[27,28] Yet the lack of data on many pollutant risks (e.g., earlier National Academy warnings about data gaps for childhood neuro-developmental effects of pesticides-herbicides)[14] militates against the disclosure and understanding conditions for informed consent. People do not receive right-to-know disclosure forms, either distributed in their neighborhoods by industries responsible for toxic releases, or available when they purchase pesticide-laden foods. They are likewise unaware, for instance, that their children are at much higher pesticide risks than adults. Consequently, public consent to imposed industrial and agricultural risks like pesticides (from which people receive far less benefit than do polluters) is much less likely than in the case of medical consent, e.g., to some drug, from which they are more likely to benefit. Because such consent is less likely, anything that increases pollutant exposure (as moving from LNT to HD would do) exacerbates ethical problems with consent and hence is ethically worse.

Can HD meet the second basic requirement of all classical codes of biomedical ethics, that subjects bearing some imposed risk have an acceptable risk-benefit ratio?[17,28] This rule requires medical experiments and societal uses of toxins to satisfy norms of distributive equity, so that most benefits of risk-imposition do not go to risk imposers, or even to society as a whole, while most risks are borne only by a subset of people. In other words, it is unethical to use some risk victims as means to the end of others, even the end of benefits for all of society. Especially it is unethical to use some risk victims as means to the end of greater benefits for risk imposers, such as pesticide-herbicide manufacturers. Yet as mentioned earlier, if the National Academy is right, then current pesticide-herbicide regulations fail to have an adequately protective risk-benefit ratio for children.[14] Because accepting HD (instead of LNT) would make children's risk-benefit ratios, at least for pesticide-herbicide responses, even worse, HD would exacerbate violations of this second key rule of biomedical ethics and thus create a worse ethical situation.

Can HD meet the third important norm of biomedical ethics, that no risk impositions, whether of medical subjects or victims of toxins like pesticides-herbicides, should result in targeting a special group of people who will bear significantly higher risks?[17,28] At least for the case of herbicides-pesticides, it is clear that their most damaging effects are borne by children. If so, weakening these already-defective, herbicide-pesticide standards (by accepting HD instead of LNT, as Calabrese proposes) would result in an even worse targeting of a vulnerable group, children, and hence would result in an ethically worse situation.[29]

If the preceding arguments are correct, HD proponents fall victim not only to logical fallacies like irrelevant thesis and confusing necessary and sufficient conditions, but also to at least 5 different sorts of ethical errors. As a consequence, HD is logically, scientifically and ethically invalid. Why is this invalidity sometimes unrecognized? Perhaps because researchers commit the fallacy of equivocation – for example, using the same term, "hormesis," to refer to three logically distinct claims, H, HG, and HD. Cook and Calabrese commit this fallacy when, under the heading "FDA Regulation of Hormesis," they refer to themselves as "proponents of hormesis" and talk about "regulation of hormesis."[4] Obviously they should have said "proponents of HD," and "regulation via HD," since H is not controversial (virtually everyone is a proponent of H), and since only invalid claim HD, not valid claim H, is specifically relevant to FDA regulation. Similar fallacies of equivocation occur, when HG-HD proponents attempt to answer critics who attack HG and HD as invalid. For instance, after Thayer et al attack HG and HD,[3] Cook and Calabrese respond to these attacks by using equivocation to defend H; they say "hormetic dose-response curves have been observed for a large number of individual agents."[4] Thus Cook and Calabrese appear to be correct, but only because they use a logical fallacy of equivocation to defend a claim, H, that is not at issue. HG and HD are at issue, but since they do not (and perhaps cannot?) defend these adequately, they mislead the reader about the nature of the argument – by focusing on H.

FIVE REFORMS TO HELP PROMOTE ACCURATE AND ETHICAL ANALYSIS OF LOW-DOSE RESPONSES

Given conflicting claims about H, HG, and HD, there are at least 5 ways in which low-dose debates and relevant research could be logically, scientifically, and ethically improved. As just suggested, the first needed improvement is (1), to distinguish claims H, HG, and HD in all research and writing, so as to avoid logical, scientific, and ethical fallacies arising from confusing three quite different claims of quite different logical and scientific validity. For instance, in their first paragraph, Cook and Calabrese say that "the concept of

hormesis….has not been without its detractors. One paper critical of the concept was published last year in this journal (Thayer et al 2005)."[4] Yet here Cook and Calabrese commit the fallacy of equivocation and confuse H, HG, and HD. Contrary to their claim, the Thayer et al paper is not critical of "the concept of hormesis," H. Rather, as is obvious from their paper, Thayer et al are critical of HG and HD.[3] As this example illustrates, often when critics challenge HG and HD, their proponents erroneously allege that the critics are challenging H. HG-HD proponents thus fail to respond to their critics' charges because they commit the logical fallacy of falsely attributing straw-man arguments (against H) to their opponents. Because straw-man arguments are far weaker than what opponents actually argue (against HG and HD), using these erroneous arguments appears to (but does not really) defeat opponents.

The fallacy of equivocation also occurs, for instance, when Cook and Calabrese say "the hormetic model also provides decision makers in regulatory agencies with a much broader array of options in the risk assessment process."[4] If this is a claim about H, it is obviously false, because H is not generalized, yet only generalized science is relevant for regulation. Likewise, if this is a claim about HG, it also is obviously false, because regulatory options require satisfaction of at least 5 democratic and ethical conditions (see earlier remarks), like informed consent, whereas HG is a purely scientific claim. Hence, the quoted remark appears to be saying that HD would theoretically provide regulatory decisionmakers with more options – a claim that requires extensive ethical support, not given by the authors, along the lines argued in the previous section. By thus equivocating by using H for HD, proponents are not obviously wrong when they fail adequately to support HD. Yet if the authors are to avoid logical fallacies, and if they mean HD, they should say HD, not the vague "hormetic model" – which could mean either H, HG, or HD.

The second improvement, in analysis of H, HG, and HD, also was defended earlier. It is (2) to treat low-dose toxins and carcinogens as pharmaceuticals, so that they might be fully tested, then regulated by the US Food and Drug Adminstration.[3] Obviously there is no reason to expose the entire US population to chemotherapeutic agents having a favorable benefit-risk ratio only for cancer patients, not most of the population. Partly because of rights to equal treatment and to self-determination, similar arguments hold for low-dose pollutants and the population-subsets they might harm or benefit.[29] Thus, without harming others, those who seek chemotherapy or low-dose-pollutant benefits can obtain them through proper individual dosage.

A third improvement needed for accurate scientific and ethical analysis of low-dose responses is (3) to encourage those who would benefit most, financially, from weakened pollution laws to fund research on H, HG, and HD. While Calabrese, Baldwin, and Cook make important points about not ignoring hormesis, responsibility and fairness dictate that those who would profit most from regulatory implementation of H, HG, and HD should either bear most of this research burden or fund independent, non-conflicted groups to do it.[29] For example, if organophosphate and related pesticides "comprise the majority of cholinesterase inhibitors that are offered by the hormesis proponents as examples of chemicals that may be beneficial at low doses," [8] chemical companies should fund the relevant research because they would profit most from HD.

A fourth improvement, needed to reliably analyze H, HG, and HD – and to follow research ethics[16,30] – is (4) to urge hormesis researchers to reveal all sources of funding, thus all potential conflicts of interest. Such revelations are especially needed, as Calabrese and Baldwin note that "the external influence of the enormous cost of environmental cleanups and the proper allocation of limited societal resources have strongly encouraged a…reexamination of… hormesis."[2] Others say something similar about chemical-industry motives regarding low-dose exposures.[31,32] A recent US National Academy of Sciences' report warned: "pesticide manufacturers" and other "economically interested third parties" are funding and conducting studies "to justify reducing" chemical-safety standards, "thereby increasing the acceptable or safe human exposure level…that might otherwise have been precluded under [current] …safety standards."[1]

Likewise, the US military, long acknowledged as the nation's worst polluter, has obvious potential (financial) conflicts of interest regarding low-dose pollutants. It is responsible for more than 15 million contaminated US acres, including 10 percent of all the worst US pollution sites (those having Superfund designation). Among more than 2300 contaminated military sites, 39 states have 130 heavily polluted military bases, all Superfund sites. One contaminant is rocket fuel, whose main ingredient is perchlorate. Especially dangerous to children's IQ, hearing, speech, and motor skills, perchlorate from military bases in Arizona, California, and Nevada, alone, contaminates the drinking water of 20 million people.[33] Yet because of costs, the US military has fought to reduce cleanup. The Pentagon wants to cut $ 4 billion per year in environmental cleanup (< 1 percent of the annual US military budget), and since 2001 the US military has failed to implement 70 federal-cleanup agreements for military bases. Yet 1 in 10 US citizens – 29 million people – live within 10 miles of military Superfund sites, and the 1986 Defense Environmental Restoration Program requires full cleanup. Many state attorneys-general, as well as city- and state-government water and waste-management agencies have sued the military, to force clean-up – which mostly has not occurred. Denver's Lowry Air Force base presents a typical case of military noncompliance with environmental-health laws. Partly because it claims low-dose pollutants are not harmful, the Air Force has refused to meet a state order to clean the 22 Lowry acres it still owns, and it has refused to reimburse the Colorado redevelopment authority for the $ 15 million cleanup that was necessary to protect homeowners from dangerous Lowry wastes left on other land by the Air Force.[34,35,36,37]

Besides the chemical industry and the military, the nuclear industry likewise has potential financial conflicts of interest regarding low-dose pollutants. President of the International Commission on Radiological Protection, Roger Clarke admitted that costs for reactor decommissioning and for radioactive waste cleanup ($ 1 trillion for US nuclear-weapons facilities alone), not science, are driving proposals to weaken low-dose radiation protection.[38]

Substantiated by US-government oversight agencies, Congress, and National Academy reports (see above), such claims suggest that chem-

ical, nuclear, and military interests all have potential conflicts of interest regarding low-dose pollutants and would gain from weakened regulations. One obvious way to address this conflict, analogous to what major environmental-health journals like *Environmental Health Perspectives* have done, is to require those who publish anything anywhere, on low-dose exposures, to reveal (in their publications, on their websites, and in their resumes) all funders of their research. Many publications of those who argue for HG and HD, for instance, are funded by groups having conflicts of interest. Calabrese acknowledges long-term US Air Force funding, [7,2,4] and Cook acknowledges consulting with Dow Chemical, a major pesticide manufacturer.[4,20] Two Calabrese reports are listed as publications of the Texas Institute for Advancement of Chemical Technology,[39,40] which is funded by Dow, BASF, Bayer, Shell Chemical, and Syngenta.[8] Such acknowledgements deserve praise, but they are incomplete. For instance, in Calabrese's online resume, 3 of 9 sources of "current research support" are not listed, yet these unlisted sources are responsible for a total of $810,000 given to Calabrese.[41]

A final ethical reform needed in H, HG, and HD research is (5) to address higher public-health and ethical priorities first.[17] Consider several facts. (A) The classic report of the US National Academy of Sciences says current pesticide regulations do not adequately protect children.[14] (B) The World Heath Organization says "only a small fraction of all childhood cancers" is associated with heredity, genetics, infections, and viruses; instead environmental pollutants appear "to play a major role," and air pollution alone is associated with up to half of all childhood cancers.[42] (C) US National Institutes of Health and National Academy of Sciences studies estimate that industrial and agricultural toxins cause about 60,000 annual US premature, fatal cancers, or about 10 percent of total cancer deaths.[43,44,29] (D) A 2002 *New England Journal of Medicine* study put the figure even higher. In its long-term study of 90,000 twins, it distinguished infection- and genetically-based, from environmental cancers, then concluded: "the overwhelming contribution to the causation of cancer in the population of twins that we studied was the environment."[45] (E) It is a public-health truism that the vast majority of potentially harmful chemicals in use – tens of thousands of them – has never been adequately tested. (F) Another public-health truism is that almost no multiple-chemical exposures, as occur in the real world, have been tested. Given the situation indicated by (A)-(F), what should be society's higher public-health priority? Should it be testing individual pollutants for low-dose beneficial effects (having little real-world applicability, given the preceding arguments)? Or should it be tracking down causes of environmental death and disease, most of which have not been adequately identified or tested? With valid arguments, HG and HD proponents might be able to make a case for the first priority. Because of their invalid arguments for HG and HD, public health easily dictates the second priority.

OBJECTIONS

In response to the preceding arguments (that although claim H is obviously true, claims HG and HD are logically, scientifically, and ethically invalid), several objections might be made. These include objections that (i) because hormesis is not defined as beneficial, it does not fall victim to some of the counterexamples given earlier; (ii) that proper allocation of societal resources argues for HG and HD; and (iii) that, contrary to earlier claims, HG and HD proponents do deal with low-dose effects on sensitive populations. Consider these objections in order

First, HG and HD proponents like Calabrese object that "beneficial/harmful [thus adaptive] effects should not be part of the definition" of hormesis.[4] However, this response is logically invalid for two main reasons. *First,* if HG-HD proponents like Calabrese contradict their earlier claims and say H does not, by definition, involve beneficial or adaptive effects,[5,6] they thereby beg the question of changing pollutant regulation because of H. Only accounts of H as adaptive or beneficial would justify the regulatory and risk-assessment changes they propose.[46] A *second* problem is that HG-HD supporters face a logical dilemma. On one hand, if they say H is not defined as beneficial or adaptive, as just noted, they beg the regulatory question. On the other hand, as argued earlier, if they say hormesis is beneficial or adaptive, they cannot generalize to HG because their claims are inconsistent with scientific evidence showing low-dose responses are often beneficial for some endpoints, but harmful for others. For example, although Calabrese and Baldwin say low-dose cadmium decreases testicular tumors in rats,[5] others report increases in prostate tumors.[46] HG and HD proponents thus have a choice between begging the question of regulatory applicability (H not defined as beneficial-adaptive), or making claims that are inconsistent with replicated scientific findings (H defined as beneficial-adaptive).

HG and HD proponents like Calabrese also object that their position is justified ethically on grounds of "proper allocation of limited societal resources,"[2] so that "the limited resources of all parties could be redirected to new agents. Control and remediation costs will be less because….resources could be redirected to other agents or…to capital investments."[20] This objection begs the question of whether cost-effectiveness arguments are ethically legitimate reasons for HG-HD. After all, one could not use cost-effectiveness to ethically justify murder-for-hire, racial discrimination, or human-rights violations, because cost-effectiveness arguments presuppose the prior ethical acceptability of the cost-cutting methods they sanction. Murder-for-hire obviously is not an ethically defensible method. But if not, HG-HD proponents must provide arguments for HD's ethical acceptability, as discussed earlier, not presuppose or beg it. Because HD proponents beg this ethical question, in appealing merely to cost-effectiveness, they presuppose the ethical validity of free-market environmentalism. This is the view that pollution ought to be controlled by the market, not regulations, and that pollution ought to be allowed whenever it is not cost-effective for polluters to reduce it.[29] But free-market environmentalism is ethically invalid because it takes no account of who causes the pollution, who benefits from it, who suffers from it, whether victims consent to it, whether it is distributed equitably and compensated, whether it results from polluter negligence or irresponsibility, and so on.

Another problem with Calabrese's cost-effectiveness objection is its committing a fallacy of aggregation. Alleging that accepting HG-HD would allow "proper allocation of limited societal resources," [2] objectors like Calabrese aggregate and call resources "societal," when they are mainly resources of polluters. If polluters are responsible for and profit from pollution, they ought to spend money to control it, money that is private and not public. In invalidly aggregating and confusing private and public (governmental) resources, these objectors invalidly allege that society will save resources in adopting HG-HD. In reality, however, the main beneficiaries would be private, polluters, while those most harmed, the greatest losers, would be the public – victims like children.

Apart from its fallacy of invalid aggregation, this second objection ignores distributive equity, fairness, and responsibility for one's actions (e.g., polluting). To see why it unethically gives polluters a free ride, at the risk of public health, consider an analogous case. Suppose someone says: " 'proper allocation of limited resources' requires setting most murderers free, since most never strike again, their threat to society is extremely low, and trial-incarceration is extremely costly." If society should never allow the costs of a murderer's prosecution and incarceration to trump ethical considerations of justice, fairness, responsibility, and compensation, society likewise should not assume that alleged polluters' costs can always trump the same ethical considerations.

A third problem with the second or cost-effectiveness objection is that HG-HD may not actually save costs overall. Obviously weakened regulations save polluter costs, but the cost-effectiveness objection begs the question (alleges, without evidence) that HG-HD implementation would save "societal resources." [2] If one counts pollution's market and non-market costs, including those to ecosystem services, individual health, work days, and so on, HG-HD likely would raise total societal costs.[47] Regardless, objectors need to empirically substantiate, not beg, the question of whether HD saves "societal resources."

A third objection from HG-HD proponents might be that their arguments do take account of sensitive groups, like children. Regarding studies of high-risk groups, Calabrese and Baldwin admitted (i) "that in about 20 percent of the cases, a hormetic response was not seen."[4,48] They also claim (ii) that if society protected these high-risk groups – by continuing to follow a LNT, rather than HD, default rule – "the general public likely could suffer an increased risk to a preventable burden of disease."[4]

Contrary to the preceding claims, (i) and (ii) do not support HG-HD, as Calabrese and others maintain. If 20 percent of cases illustrate no hormetic effect, then LNT, not HD, better protects this 20 percent. But if so, this argues against HD because it contradicts two major claims of HD proponents. *First*, LNT's superior protection of the 20-percent-high-risk group (claim i) contradicts HG, the claim (on which HD relies) that H is "generalizable across biological model, endpoint measured, and chemical class."[7] *Second*, LNT's superior protection of this 20 percent (claim i) also contradicts allegations that "the hormesis model clearly outperforms" either T or LNT models.[49] Thus, if objectors' claim (i) is true, it follows that it has refuted two of the objectors' own HG-HD arguments.

What about claim (ii), that if society protected these high-risk groups – by continuing to follow a LNT rather than HD default rule – "the general public likely could suffer an increased risk to a preventable burden of disease"[4]? Here the objectors provide no empirical documentation, whatsoever, to support claim (ii). Thus they again beg the question. Moreover, claim (ii) also is highly implausible, given the earlier arguments that additive and synergistic effects of multiple exposures together yield exposures that are no longer low-dose. By begging the question of (ii) and ignoring empirical data on total doses and synergistic effects, HG and HD proponents again appear to be doing ideology, not science and not ethics.

CONCLUSION

Analysis of low-dose pollution effects is important and ought not be ignored. If it is to be accomplished with logical, scientific, and ethical rigor, however, at least 5 reforms are needed. These include (1) avoiding logical fallacies like equivocation by distinguishing claims H, HG, HD, rather than using only "hormesis" labels; (2) protecting informed consent by assessing and regulating low-dose exposures as pharmaceuticals; (3) ensuring fairness and responsibility by having those, who would profit most financially, pay for H, HD, and HG research; (4) following research ethics by having researchers' reveal all potential conflicts of interest; and (5) protecting rights to equal protection by first pursing research that is more important to public-health priorities. Since many researchers do not follow (1)-(5), they bear the burden of proof to defend both their ethics and their reasons for not accomplishing (1)-(5).

REFERENCES

1. National Research Council. Intentional Human Dosing Studies for EPA Regulatory Purposes. National Academy Press: Washington, DC, 2004, pp. 110, 112-113, 3.

2. Calabrese EJ, Baldwin EJ. Hormesis: U-Shaped Dose Responses and Their Centrality in Toxicology. Trends in Pharmacological Sciences 2001; 22: 286, 290, 285 of 285-291.

3. Thayer KA, Melnick R, Burns K, Davis D, Huff J. Fundamental Flaws of Hormesis for Public Health Decisions. Environmental Health Perspectives 2005; 113:1272-5.

4. Cook R, Calabrese EJ. The Importance of Hormesis to Public Health. Environmental Health Perspectives 2006; 114: 1631-1635; online 7-10-06 at http://www.ehponline.org/members/2006/8606/8606.pdf and accessed 11-22-06.

5. Calabrese EJ, Baldwin LA. Hormesis: The Dose-Response Revolution. Annual Review of Pharmacology and Toxicology 2003; 43: 191 of 175-97.

6. Calabrese EJ, Baldwin LA. Defining Hormesis. Human and Experimental Toxicology 2002; 21: 91 of 91-97.

7. Calabrese EJ. Historical Blunders: How Toxicology Got the Dose-Response Relationship Half Right. Cellular and Molecular Biology 2005; 51: 643 of 643-654.

8. Axelrod D, Burns K, Davis D, Von Larebeke N. Hormesis – An Inappropriate Extrapolation from the Specific to the Universal. International Journal of Occupational and Environmental Health 2004; 10: 336-8 of 335-339.

9. Scientific Advisory Board and FIFRA Scientific Advisory Panel. Comments on the Use of Data from Testing of Human Subjects, EPA-SAB-EC-00-017. EPA: Washington, DC, 2000.

10. Kociba RJ, Keyes DG, Beyer JE. Results of a Two-Year Chronic Toxicity and Oncogenicity Study of 2,3,7,8-Tetrachlorodibenzo-p-dioxin in Rats. Toxicology and Applied Pharmacology 1978; 46: 279-303.

11. Mead MN. Sour Finding on Popular Sweetener. Environmental Health Perspectives 2006; 114, online at http://www.ehponline.org/docs/2006/114-3/ss.html and accessed 11-22-06.

12. Reir SE, Martin DC, Bowman RE, Domowski WP, Becker JL. Endometriosis in Rhesus Monkeys (Macaca mulatta) Following Chronic Exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Fundamentals of Applied Toxicology 1993; 421: 433-41.

13. Calabrese EJ, Baldwin LA. Toxicology Rethinks Its Central Belief. Nature 2003; 421: 691 of 691-2.

14. US National Research Council. Pesticides in the Diets of Infants and Children. National Academy Press: Washington, DC, 1993.

15. Shrader-Frechette KS. Risk and Rationality. University of California Press: Berkeley, 1991, pp. 100-145.

16. Shrader-Frechette KS. Ethics of Scientific Research. Rowman and Littlefield: New York, 1994, pp. 9-12, 23-118.

17. Association of Schools of Public Health. Ethics and Public Health, edited by Bruce Jennings, Jeffrey Kahn, Anna Mastroianni, Lisa Parker. ASPH: Washington, DC, 2003.

18. Hume D, A Treatise of Human Nature. Clarendon Press: Oxford, 1975.

19. Aristotle, Nicomachean Ethics, trans. Terence Irwin. Hackett: Indianapolis, 1985.

20. Calabrese EJ, and Cook R. Hormesis: How It Could Affect the Risk Assessment Process. BELLE 2005; 12: 26 of 22-27.

21. Calabrese EJ, Baldwin LA. Reevaluation of the Fundamental Dose: Response Relationship. BioScience 1999; 49: 725-732.

22. Lang L. Strange Brew: Assessing Risk of Chemical Mixtures. Environmental Health Perspectives 1995;103: online at http://www.ehponline.org/docs/1995/103-2/focus.html and accessed 11-22-06.

23. Calabrese EJ, Baldwin LA. Radiation Hormesis. Human and Experimental Toxicology 2000; 19: 2.

24. Calabrese EJ, Baldwin LA. Scientific Foundations of Hormesis. Critical Reviews of Toxicology 2001: 31: 351-624.

25. National Research Council. Health Risks from Exposure to Low Levels of Ionizing Radiation, BEIR VII. National Academy Press: Washington, DC, 2005.

26. Shrader-Frechette KS. Using A Thought Experiment to Clarify a Radiobiological Controversy. Synthese 2001; 128: 319-342.

27. Faden R, Beauchamp T. A History and Theory of Informed Consent. Oxford University Press: New York, 1986.

28. Beauchamp T, Childress J. Principles of Biomedical Ethics. Oxford University Press: New York, 1989.

29. Shrader-Frechette KS, Taking Action, Saving Lives: Our Duties to Protect Environmental and Public Health. Oxford University Press: New York, 2007.

30. Krimsky S, Science in the Private Interest. Rowman and Littlefield: New York, 2003.

31. Oleskey C, Fleishman A, Goldman LR. Pesticide Testing in Humans. Environmental Health Perspectives 2004; 112: 114-119.

32. Lockwood, AH. Human Testing of Pesticides. American Journal of Public Health 2004; 94: 1908-1915.

33. US National Research Council. Health Implications of Perchlorate Ingestion. National Academy Press: Washington, DC, 2005.

34. US Congress, House of Representatives, Committee on Energy and Commerce. Department of Defense Attempts to Exempt Itself from Environmental Laws and Regulations. CEC: Washington, DC, 2006, available at http://www.house.gov/commerce_democrats/DODexemptions/dod.shtml and accessed 11-22-06.

35. US General Accounting Office. Military Munitions: DOD Needs to Develop a Comprehensive Approach for Cleaning Up Contaminated Sites, GAO-04-147. GAO: Washington, DC, 2003.

36. Environmental Working Group. Rocket Fuel in Drinking Water: Perchlorate Pollution Spreading Nationwide. EWG: Washington, DC, 2006), available at http://www.ewg.org/reports/rocketwater/ and accessed 11-22-06.

37. Eisler P. Pollution Cleanups Pit Pentagon Against Regulators. USA Today 10-14-04, available at http://www.usatoday.com/news/nation/2004-10-14-cover-pollution_x.htm and accessed 11-22-06.

38. Clark R. Control of Low-Level Radiation Exposure: Time for a Change? Journal of Radiological Protection 1999; 19: 107-115.

39. Calabrese EJ, Baldwin LA. Chemical Hormesis: Scientific Foundations, unpublished final report of the Texas Institute for Advancement of Chemical Technology, 1998, online at http://cheweb.tamu.edu/tiact/index_files/Page536.htm and accessed 11-22-06.

40. Calabrese EJ, Baldwin LA, Holland C, Hormesis: A Highly Generalizable and Reproducible Phenomenon with Important Implications for Risk Assessment. Risk Analysis 1999; 19: 261-281.

41. Calabrese EJ. Faculty. University of Massachusetts Intercampus Graduate School of Marine Sciences and Technology, 2006, available at http:/www.umassmarine.net/faculty/showprofs.cfm?prof_ID=30 and accessed 11-22-06.This was the most up-to-date resume found on the internet for Dr. Calabrese, and it was copyrighted 2006, by his employer, as an "official page/publication of the University of Massachusetts. [copyright form symbol insert here] 2006"; this same resume (with the same material), accessed January 30, 2007, appears again as "an official page/publication of the University of Massachusetts. [copyright form symbol insert here] 2007." Although an older, more complete resume ("Edward J. Calabrese, Ph.D., Curriculum Vitae, April 2002") appears on the University of Massachusetts website, with a 2002 date, at http://people.umass.edu/nrephc/EJCCVApril02.pdf (accessed 30 January 2007), it has more than $3 million in grants/contracts whose sources of funding are not named; some named sources include Exxon, ARCO, Dow Chemical, Reynolds, Rohm and Haas. Even an out-of-date or shortened resume should not selectively reveal funding sources, both because readers will be mislead, because readers expect that state-university employees, in particular, reveal their funding sources, and because most science-ethics groups endorse full disclosure.

42. World Health Organization, Effects of Air Pollution on Children's Health. WHO*:* Bonn, 2005, pp. 144, 155.

43. Health and Human Services and National Cancer Institute. Health Status Objectives. Cancer 1991; 16: 416-440.

44. McGinnis JM. Attributable Risk in Practice. In Institute of Medicine. Estimating the Contributions of Lifestyle-Related Factors to Preventable Death. National Academy Press: Washington, DC, 2005, pp. 17-19.

45. Lichtenstein P, Holm N, Verkasalo P, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthee A, Hemminki K, Environmental and Heritable Factors in the Causation of Cancer. New England Journal of Medicine 2002: 343: 78-85.

46. Thayer KA, Melnick R, Huff J, Burns K, Davis D. Hormesis: A New Religion? Environmental Health Perspectives 2006; 114: 632-633.

47. US National Research Council, Valuing Ecosystem Services. National Academy Press: Washington, DC, 2004.

48. Calabrese EJ, Baldwin LA. Hormesis and High-Risk Groups. Regulatory Toxicology and Pharmacology 2002; 35: 414-428.

49. Calabrese EJ, Baldwin LA. Hormesis: A Revolution in Toxicology, Risk Assessment and Medicine. EMBO Reports 2004; Supplement 1:S37-40.

HORMESIS, ETHICS, AND PUBLIC POLICY: AN OVERVIEW

Kevin C. Elliott
Department of Philosophy
University of South Carolina
Columbia, SC 29208
Phone: (803) 777-3735
Fax: (803) 777-9178
Email: ke@sc.edu

I would first like to thank Ed Calabrese for inviting me to be the guest editor for this journal issue and for discussing these issues with me. The first step toward formulating ethical public policies in response to the hormesis phenomenon is to have an open, transparent discussion of the major issues involved, so I applaud Ed for his commitment to advance thoughtful discussion of these topics. This essay provides an overview and synthesis of the major points made in the articles submitted for this issue. I have organized the questions and concerns raised in these contributed pieces into three general categories: (1) scientific issues, (2) practical concerns, and (3) "explicitly ethical" considerations. I refer to the third category as "explicitly" ethical, because all three categories involve ethical considerations in at least a broad sense. Nevertheless, the third group of questions involves issues that are ethical in a particularly obvious and traditional sense, such as balancing risks and benefits, obtaining informed consent, and promoting distributive justice. After discussing each of the three categories in turn, I will summarize some of the suggestions provided by these essays for responding to hormesis in the future.

SCIENTIFIC CONSIDERATIONS

The essays by Shrader-Frechette and by Mayo and Spanos both emphasize that many aspects of scientific methodology and research choices are ethical in at least a broad sense, insofar as they influence policy decisions that affect the public. Under this category of ethically-relevant scientific considerations one might include issues such as the evaluation of controversial scientific inferences and the need for further research in particular policy-relevant areas. Perhaps the most common scientific concern raised in the preceding articles is that the generalizability of hormesis may not be adequately established for the purposes of formulating public policy (see e.g., Douglas, Mayo and Spanos, Shrader-Frechette). Mayo and Spanos emphasize that there are a number of statistical pitfalls associated with searching for evidence of hormesis in existing studies. Shrader-Frechette worries that proponents of applying hormesis to public policy have not conducted epidemiological field studies, which would be most likely to mimic real-world exposure conditions. Thus, she suggests that these proponents may be making a similar sort of inductive error that opponents of hormesis made in the past (namely, overgeneralizing a particular set of dose-response relationships without adequate evidence). Douglas argues that, although the hormesis hypothesis may have significant fertility and explanatory power on its side, the history of science is replete with cases of fertile and explanatory theories that have nevertheless turned out to be false.

Hoffmann and Stempsey emphasize that hormesis proponents have been taking steps to respond to these concerns about generalizability (see also Cook and Calabrese 2006). Nevertheless, many ethicists demand fairly high standards of evidence for policy decisions that could harm the public if they turn out to be wrong, especially when there might be alternative policies that pose less serious risks (see Hansson). The contributed essays make a number of suggestions that could be helpful for scientists who attempt to meet these high standards of evidence in the future. First, Shrader-Frechette argues that researchers should distinguish carefully in their writings between three different claims: (a) that hormesis occurs, (b) that hormesis is generalizable across biological model, endpoints measured, and chemical class, and (c) that hormesis should be the default model in risk assessment and management. One might add that it would be helpful to be more precise about what it means to say that hormesis is "generalizable." Does it mean that, for each biological model, endpoint, and chemical class, there is at least one example of a hormetic dose-response relationship? Or does it mean that a particular percentage (say, 50%) of toxic chemicals exhibit hormetic dose-response relationships (on at least some endpoints, in at least some biological models)? Or does it mean that, if one were to formulate a comprehensive list of the dose-response relationships for every toxic chemical on every endpoint in every biological model, some percentage of those relationships (say, 50%) would be hormetic? The contributed essays suggest that scientists might alleviate part of the ethical controversy over hormesis by making and testing more precise claims about the generalizability of hormesis.

Douglas makes another suggestion for those studying hormesis, namely, that they would do well to formulate and test predictions about the precise conditions under which hormetic dose-response relationships are likely to occur. Along the same lines, Mayo and Spanos argue that hormesis researchers should consider at least three improvements or alternatives to their previous literature studies in order to defuse statistical concerns about their methodologies: (a) obtaining reliable estimates of the control incidence rates for alleged hormetic effects, (b) examining (rather than ignoring) cases that have low or zero disease incidence in controls, and (c) conducting new, genuinely controlled studies of hormetic effects with several doses in the hormetic range.

Although concerns about the generalizability of hormetic dose-response relationships were especially prominent in the submitted essays, they discuss several additional scientific issues. For example, Douglas and Shrader-Frechette argue that it is important to determine when hormetic dose-response relationships are likely to translate into genuine health benefits, because some hormetic effects on individual endpoints may be harmful from the perspective of the organ-

ism as a whole. Elliott (2006b) has previously suggested along these same lines that it may be necessary to develop sharper criteria for the sorts of effects that count as genuine benefits and harms from the perspective of the organism. Moreover, Shrader-Frechette emphasizes that seemingly beneficial short-term hormetic effects may be harmful over the long term (see also Elliott 2006b). Proponents of hormesis have already been seeking evidence that could address these issues (see Cook and Calabrese 2006), but the contributed essays emphasize that more scientific information about these questions is likely needed before policy makers would incorporate hormesis as a default assumption in risk assessment.

PRACTICAL CONSIDERATIONS

For the purposes of this essay, I refer to "practical" considerations as factors associated with the social, institutional, and environmental context that influence what sorts of public policies are ethically appropriate. Many of the submitted essays distinguish the acceptance of hormesis as a biological hypothesis from the acceptance of hormesis as a basis for making policy (see e.g., Douglas, Hansson, Hoffmann and Stempsey, Shrader-Frechette). They emphasize this distinction for at least two reasons: (1) the standard of proof appropriate in the scientific context may be different from what is demanded in the policy arena; and (2) there may be "practical" reasons that, even if hormesis occurs, it would not be feasible to act on it. The essays discuss several of these practical reasons.

The most prominent concern is that it may be too difficult to expose people to levels of toxins that are actually in the hormetic zone. Sandin calls this the "fine-tuning" problem, and Shrader-Frechette similarly highlights the challenge of "titrating" beneficial levels of chemicals for individuals (see also Elliott 2006b). This problem is complicated by the fact that chemicals may have aggregative or synergistic effects, which could already be exposing many people (and especially children or other sensitive populations) to dose levels that are no longer beneficial (see Hansson, Sandin, Shrader-Frechette). The articles suggest some avenues for addressing this concern. First, Hansson argues that more effort should be put into developing exposure assessments, which are often a "weak link" in risk assessments. Second, several authors propose that one could allow individuals to choose the levels of hormetic chemicals to which they are exposed by regulating them as pharmaceuticals or "vitamin" supplements (see Douglas, Elliott 2006b, and Shrader-Frechette).

Some of the submitted articles provide cautions, however, against dismissing policy implications of hormesis too quickly just because of practical concerns. For example, Hoffmann and Stempsey claim that public health is ultimately more likely to be served by using the most scientifically accurate dose-response models (although they acknowledge that there is still a good deal of uncertainty surrounding the hormetic model). Moreover, Renn suggests that, if one were to assume that there is adequate scientific evidence for hormesis, it would be ethically wrong not to consider it in some fashion in the regulatory balancing of risks and benefits. Other contributed essays suggest at least two potential responses to these claims on behalf of hormesis. First, even if the hormetic model describes the effects of chemicals correctly when individuals are exposed to one toxin at a time (and under precisely controlled conditions), the threshold or linear models may actually provide a better estimate of the effects of toxins when people are exposed to several at a time under real-life conditions (see Shrader-Frechette). Second, Sandin notes that, in the real world, there can be significant practical benefits from formulating policies that are relatively simple, such as the ALARA (as low as reasonably achievable) principle for radiation policy. The contributed essays do not settle these issues but rather highlight the need for further discussion and investigation.

"EXPLICITLY ETHICAL" CONSIDERATIONS

Hansson, Renn, and Shrader-Frechette all emphasize that the ethics of risk assessment and management cannot be reduced solely to an analysis of the total balance of risks and benefits across society. One must consider such issues as which individuals are being exposed to which risks and which benefits, whether conditions for informed consent to risks are met, and whether appropriate compensation for risks is possible. The most common "explicitly ethical" issue discussed in the contributed essays is that hormesis raises questions about whether regulatory agencies should be seeking *benefits* for the population or merely trying to prevent *harms*. Multiple authors highlighted two broad approaches that regulators might take (see e.g., Douglas, Hansson, and Sandin). A broadly utilitarian approach would be to maximize the ratio of benefits to harms for the population as a whole. The other, more "deontological," approach would be to focus on individual members of the population, making sure that no one is exposed to an unacceptable level of risk without appropriate consent and compensation.

Hansson provides an intriguing summary of how these two different approaches have been accepted in different social and institutional contexts. There appear to be a variety of reasons, though, that a more individual-focused, deontological approach is likely to dominate discussions of hormesis. Among these considerations are the fact that ethicists have generally placed a higher priority on preventing harm rather than providing benefits (Hoffmann and Stempsey) and the fact that the benefits associated with hormesis appear to be relatively small compared to their potential for harm (see Douglas, Hoffmann and Stempsey, and Sandin). Several authors noted that controversies regarding the fluoridation of water illustrate the public's lack of enthusiasm for government policies that focus on maximizing benefits.

Proponents of making hormesis the default model in risk assessment have tried to escape these difficulties by suggesting that perhaps there are low-dose levels at which toxins could be beneficial for *everyone*, including sensitive sub-populations (see Cook and Calabrese 2006). Nevertheless, even if this "ideal" scenario turned out to be the case, several contributed essays emphasize that one must consider the possibility that some individuals are already receiving chemical exposures that are outside the hormetic range (see especially Shrader-Frechette). Thus, given the "deontological" flavor of much ethical thought

regarding pollution regulation, further information about the synergistic effects of chemicals and the current exposures of sensitive individuals to toxic chemicals will probably be needed in order to make a convincing case for easing regulatory policies in response to hormesis. In the meantime, it appears that the proponents of hormesis are likely to be on their strongest ethical footing when they emphasize applications in which the consideration of hormesis could clearly prevent harms to specific individuals. For example, using hormetic dose-response curves to prevent improper dosing of chemotherapeutic agents (see Calabrese and Baldwin 2001) is likely to garner much more universal enthusiasm in the near future than the application of hormesis to pollution regulation.

The essays (especially Douglas, Renn, and Shrader-Frechette) mention a number of other ethical considerations (often in the context of arguing that a narrowly "utilitarian" approach to regulation would be problematic) that may be helpful for policy makers to consider in the future. These include: (1) concerns about justice and fairness, given that disadvantaged groups and children might bear a greater proportion of the risks associated with easing chemical regulations, (2) worries about whether those exposed to risks associated with hormesis could (or would) provide informed consent to them, (3) concerns about the difficulty of compensating any individuals who might end up being harmed, and (4) issues of sustainability, political and legal implementability, and public acceptance. In response to these concerns, Renn encourages the development of analytic-deliberative processes that could provide more precise information about the values and concerns of different segments of the public (see also Elliott 2006a and 2006b; NRC 1996). In addition to advancing public consent to future policies regarding hormesis, these processes could help regulators determine whether the public would indeed be averse to "utilitarian" schemes for balancing risks and benefits

SUGGESTIONS

In the process of raising ethical considerations and concerns, the contributed essays suggest a number of promising strategies for researchers and policy makers to pursue in the future. First, scientists would do well to continue addressing scientific concerns about the generalizability of hormesis. And, at the very least, the essays highlight the importance of being clear about the precise claims that researchers are making and testing when they refer to the "generalizability" of hormesis. Second, it would be helpful to clarify both what counts as a genuine, long-term health benefit and how often hormetic effects are genuinely beneficial in this sense. Third, it seems likely that hormesis will be resisted as a default model in risk assessment until more information is available about the toxic exposures that the public is already receiving and the synergistic effects of those exposures.

Fourth, ethicists, policy makers, and the public at large will need to do some careful thinking about whether the goals of toxic chemical regulation should involve, on one hand, maximizing the ratio of benefits to risks for the population as a whole or, on the other hand, protecting individuals from health risks to which they do not consent and for which they cannot easily be compensated. As Renn emphasizes in his contribution, there are innovative analytic-deliberative mechanisms that could provide avenues for gathering both public and expert input regarding these questions (see also Beierle 2002; Kleinman 2000). Meanwhile, proponents of applying hormesis in practical contexts might do well to focus, in the near term, on cases in which the phenomenon could prevent harm to specific individuals (e.g., in the medical setting). These applications would showcase the clearest strengths of hormesis from an ethical perspective while debates about the ethical ramifications of hormesis in the realm of pollution regulation continue to play out.

REFERENCES

Beierle, T. (2002), "The Quality of Stakeholder-Based Decisions," *Risk Analysis* 22: 739-749.

Calabrese, E. and L. Baldwin (2001), "Hormesis: U-Shaped Dose Responses and Their Centrality in Toxicology," *TRENDS in Pharmacological Sciences* 22: 285-291.

Cook, R. and E. Calabrese (2006), "The Importance of Hormesis to Public Health," *Environmental Health Perspectives* 114: 1631-1635.

Elliott, K. (2006a), "The Case of Chemical Hormesis: How Scientific Anomaly Shapes Environmental Science and Policy," in D. Guston and D. Sarewitz (eds.) *Shaping Science and Technology Policy: The Next Generation of Research*. Madison, WI: University of Wisconsin Press.

Elliott, K. (2006b), "Hormesis and Environmental Policy: An Ethical Analysis," *Public Affairs Quarterly* 20: 31-53.

Kleinman, D. (2000). *Science, Technology, and Democracy*. Albany, NY: State University of New York Press.

National Research Council (1996), *Understanding Risk: Informing Decisions in a Democratic Society*. Washington, D.C.: National Academy Press.

The 7th Annual International Conference on

DOSE-RESPONSE 2008:
Implications for Toxicology, Medicine and Risk Assessment

April 29-30, 2008

The Annual Meeting of the International Dose-Response Society

Dose-Response Conference

Please visit our website for more information, Abstract Submission Guidelines and Abstract Submission

Visit us on the web: ***www.Dose-Response.org***