HENK VAN DEN BELT and BART GREMMEN

BETWEEN PRECAUTIONARY PRINCIPLE AND "SOUND SCIENCE": DISTRIBUTING THE BURDENS OF PROOF

(Accepted June 1, 2001)

ABSTRACT. Opponents of biotechnology often invoke the Precautionary Principle to advance their cause, whereas biotech enthusiasts prefer to appeal to "sound science." Public authorities are still groping for a useful definition. A crucial issue in this debate is the distribution of the burden of proof among the parties favoring and opposing certain technological developments. Indeed, the debate on the significance and scope of the Precautionary Principle can be fruitfully re-framed as a debate on the proper division of burdens of proof. In this article, we attempt to arrive at a more refined way of thinking about this problem in order to escape from the existing polarization of views between "guilty until proven innocent" and "innocent until proven guilty." This way of thinking also enables a critical review of current demarcations between risk assessment and risk management, or science and politics, and of the morally laden controversy on the relative importance of type-I and type-II errors in statistical testing.

KEY WORDS: biotechnology, burden of proof, Precautionary Principle, type-I and type-II errors

INTRODUCTION

Ever since its insertion into the 1992 Rio Declaration on Environment and Development and a number of multilateral treaties like the 1992 Convention on Biological Diversity and the 1992 Framework Convention on Climate Change, the so-called Precautionary Principle has been a bone of contention. It is often invoked by environmentalists and opponents of biotechnology as a powerful argument to advance their cause, but just as often it is strongly resisted by advocates of free enterprise and biotech enthusiasts as an arbitrary departure from "sound science." Caught in the middle, public authorities are still groping for a plausible and acceptable interpretation of this principle, which until now has escaped a strict definition. Intuitively, the intended meaning of the Precautionary Principle seems clear enough, but its precise articulation runs up against various intricacies and subtleties. As the principle has recently been invoked to justify bans on the imports of certain food and agricultural products into

Journal of Agricultural and Environmental Ethics **15:** 103–122, 2002.

the European Union, its significance and legal force are also at stake in the ensuing trade disputes before the forum of the World Trade Organization.

An example of an (in our view) unsuccessful struggle with the intricacies of the Precautionary Principle is provided by the position paper that the Commission of the European Communities prepared on this subject in response to demands of the Council and the European Parliament (European Commission, 2000). The Commission conceded that there was no generally accepted definition of the Precautionary Principle and deliberately declined to define the term; instead, its precise meaning had to be fleshed out by decision-makers and courts of law. Nonetheless, the Commission was quite sanguine about the prospects of getting an ever-better handle on the principle, thanks to increasing practical experience of Community authorities and judicial review. So, in the Commission's view, "it would be wrong to conclude that the absence of a definition has to lead to legal uncertainty." Gregory Conko, director of food safety policy with the CEI (Competitive Enterprise Institute) in Washington, DC, severely criticizes this cavalier attitude and compares it to the famous remark from Justice Potter Stewart that he couldn't define pornography, but that he knew it when he saw it: "Leaving an innovator's (or anyone else's) legal rights undefined makes us captive to the wholly subjective judgment of politicians. In effect, regulators are given carte blanche to decide what is 'unsafe' and what is 'safe enough,' with no checks or balances to ensure that their decisions actually reduce overall risk" (Conko, 2000). Even if we allow for Conko's unabashed anti-regulation (or minimal regulation) stance, we cannot but concur with his view that the absence of an authoritative definition of the central principle will inevitably give rise to legal uncertainty.

In this article, however, we will not come up with the eagerly looked-for "definitive" definition of the Precautionary Principle that promises to solve all outstanding problems and to end all legal uncertainty. What we will do is attempt to clarify the underlying issues by consistently reconceptualizing them in terms of the proper division of the burden (or rather burdens) of proof. It is sometimes stated that adoption of the Precautionary Principle entails abandonment of the old maxim "innocent until proven guilty" in favor of the new maxim "guilty until proven innocent." Although the suggested contrast is far too crude, in our view, it rightly points to the problem of the distribution of burdens of proof as the central issue in the controversies and polemics surrounding the invocation and application of the said principle. Following this argumentation line allows us to reject, on the one hand, the absolutist interpretation of the Precautionary Principle propounded by many environmentalist organizations, and, on the

other, its blanket dismissal by free enterprise lobbyists and biotech enthusiasts in favor of a single-minded appeal to so-called "sound science." The former position is unacceptable because it involves an unfair and unreasonable distribution of burdens of proof. The latter position must also be rejected because it fails to recognize that regulatory science can work only within the framework of an ethically defensible and socially acceptable distribution of burdens of proof.

THE CASE AGAINST THE PRECAUTIONARY PRINCIPLE

In 1992, in the context of dealing with environmental hazards, the Rio Declaration presented the following formulation of the Precautionary Principle: "Where there are threats of serious or irreversible damage, lack of full scientific certainty shall not be used as a reason for postponing cost-effective measures to prevent environmental degradation." This formulation may sound unexceptional, but neither is it of much help as an effective basis for policy, as long as the meanings of the key-terms "serious," "irreversible," "damage," and "cost-effective" are not sufficiently specified and it is not decided how much scientific certainty (short of full "scientific certainty") we need before we may (or should) undertake preventive action. While the principle's scope has meanwhile been extended so as to evolve into "a key tenet of risk management in the areas of environmental protection and the protection of human, animal and plant health" (as in the European Commission's position paper of February 2000), most of the descriptions now on offer formulating its significance and import suffer from similar ambiguities.

These ambiguities have not prevented environmentalist organizations such as Greenpeace or Friends of the Earth from turning the Precautionary Principle into a staple "argument" of their campaigns. It seems to be tailor-made for that role. Precisely because the current formulations of the principle do not specify a minimal threshold or level of scientific certainty warranting preventive action, these organizations feel that even the slightest indications that new technologies might cause some possible damage for the environment or for human, animal, or plant health entirely justify the demand to halt them. If companies nonetheless persist in their wish to continue developing such (at least in theory) potentially harmful technologies, the environmentalist organizations argue, it is up to them to prove beyond any doubt that these technologies are really safe. The campaign of Greenpeace against genetically modified crops illustrates that the Precautionary Principle may indeed be invoked already on the slightest grounds. When in 1999 the scientific journal *Nature* published a report

on a laboratory experiment indicating that pollen from genetically modified Bt-maize could kill the larvae of monarch butterflies, the article was immediately taken up by Greenpeace to support its demand for a ban on GM crops (Losey et al., 1999, p. 214). In a media-oriented action, members of Greenpeace dressed as butterflies confronted a meeting of EU environment ministers held in Luxembourg in June 1999 and carried banners demanding "Give butterflies a chance." However, Greenpeace decided to ignore the press release prepared by the authors of the article in *Nature*, who declared that the results of their study, which had been conducted under rather special and artificial laboratory circumstances, could not be easily extrapolated to natural field conditions. Thus the study did not show that monarch butterfly populations in the wild were actually endangered by Bt-maize or genetically engineered crops in general.[1] (The campaign of Greenpeace against GM crops met with some success; the EU decided to freeze the approval process.)

Although environmental organizations like Greenpeace are still considered highly trustworthy by the general public (much more so than governments and private companies), critics have cast some doubt on the purity of their motives. It is argued that they are part of a larger "danger establishment," consisting of scientists, journalists, politicians, bureaucrats, and environmental organizations with a vested interest in exaggerating dangers. Greenpeace, in particular, can be likened to a multinational that is compelled by the necessity of covering its salary and operating costs to select campaign issues that will heighten concern and thus help in raising funds (Kellow, 1999). We would indeed be naïve to presume that Greenpeace and other environmental organizations only have the public good at heart.

In view of the fact that the Precautionary Principle is the mainstay in the environmentalists' case against GM crops, it should not be surprising that biotech lobby groups and advocates of free enterprise generally find fault with it, or at least with the absolutist version the principle has assumed in the hands of many environmentalists. The critics have put the finger on the sore spot. "In practice," Henry I. Miller and Gregory Conko write, "the precautionary principle establishes a lopsided decision-making process that is inherently biased against change and therefore against innovation" (Miller and Conko, 2000). The reason is that it leads its protagonists to focus mainly on the possibility that new technologies may pose theoretical risks, always hedging against the worst possible outcomes by assuming

[1] More recent field research performed in the American Mid-West seems to indicate that monarch butterfly populations are hardly affected, if at all, by the large-scale cultivation of Bt maize in this region. See Schmickle, 2000.

worst case scenarios, while ignoring the potential benefits of these same technologies or the real existing risks that could be mitigated or eliminated by them. Already in 1995 the late Aaron Wildavsky railed against the illusionary belief that by adhering to the Precautionary Principle something very valuable, to wit human or environmental health, could be got at virtually no cost whatsoever, the facile assumption being that the proposed bans and regulations themselves would have no adverse health effects (Wildavsky, 1995, esp. p. 428). Miller and Conko imagine the principle to have been applied in a counterfactual scenario to make their point: "If the precautionary principle had been applied decades ago to innovations such as polio vaccines and antibiotics, regulators might have prevented occasionally serious, and sometimes fatal, side effects by delaying or denying approval of those products, but that precaution would have come at the expense of millions of lives lost to infectious diseases" (Miller and Conko, 2000).

A point of elementary wisdom deserves to be emphasized here. The maxims "Safety first!" and "Better safe than sorry" do not point the way to a safe, risk-free world, for such a world is not to be had. Or, as some humorist once said, the only way to avoid accidents is to stay in bed, but even that would not prevent one from the risk of falling out. A distinction that was first introduced in the philosophical debate on skepticism may also be used here: the distinction between *misfortunes of the first kind* ("We reject something that, as it turns out, we should have accepted") and *misfortunes of the second kind* ("We accept something that, as it turns out, we should have rejected.") As Nicholas Rescher has shown, a would-be "risk avoider" will suffer few misfortunes of the second kind, but only at the cost of incurring relatively many of the first.[2] The old cliché "Better safe than sorry" needs at least to be balanced by another old cliché: "Nothing ventured, nothing gained."

The application of the Precautionary Principle also tends to impose an impossible burden of proof on the proponents of new technologies. In the name of absolute safety they are asked nothing less than to demonstrate conclusively that the new technologies they advocate offer no possible harm. This is a formidable, perhaps even a logically impossible task. Wildavsky repeats an observation made by Harvey Brooks to the effect that "the only proof of a negative [such as 'no possible harm'] is an impossibility theorem demonstrating that the contemplated action or reaction is contrary to the laws of nature" (Wildavsky, 1995, p. 430). Aynsley Kellow

[2] Rescher, 1995, p. 84. Rescher's distinction between "misfortunes of the first kind" and "misfortunes of the second kind" is reminiscent of the distinction between "type-I errors" and "type-II errors" in statistical reasoning.

also makes a graphical comparison: "Demanding that a negative be proved is the logical equivalent of asking people to prove that they are not witches" (Kellow, 1999, p. 6). Clearly, a game that allows one party to play by the rule "heads I win, tails you lose" is not entirely fair.

So far we have endorsed the two main points of criticism that advocates of free enterprise and biotech enthusiasts have mounted against the Precautionary Principle, at least as it is interpreted by environmental organizations. Firstly, it is used to examine only one side of the risk equation, the risk of using new technologies, and not the other side: their benefits or the "opportunity costs" in the form of foregone possibilities of mitigating known risks. Secondly, the principle tends to put a well-nigh impossible onus of proof on the shoulders of the proponents of new technologies. It all boils down to the fact that the Precautionary Principle in its absolutist interpretation precludes any balancing between the risks of harm with the potential benefits of innovation. We think that to preclude such balancing as a matter of principle is highly unreasonable. That is not to say that the risks and benefits of new technologies are all on an equal par and that their balancing is a simple and straightforward technical task that can be safely delegated to the practitioners of CBA (cost-benefit analysis), RBA (risk-benefit analysis) and kindred methodologies. Indeed, there may be good ethical reasons to assign unequal weight to various considerations and concerns. We hope to show below that the conceptualization of the problem as one of the proper distribution of burdens of proof will be helpful in this regard.

Our position can be compared to the view expressed in the report *Genetically modified crops: the ethical and social issues*, published in May 1999 by the British Nuffield Council on Bioethics (see Nuffield, 1999). The Working Party on GM crops that prepared the report noted that "on a stringent definition of the precautionary principle there could be no balancing of the risk of harms with the benefits of innovation, since even a suspicion of possible harm, no matter how ill-founded, would be sufficient to prohibit a new technology" (see Nuffield, 1999, section 7.13 and 1.14). The Working Party held that the principle in this stringent form is implausible. Nevertheless, it endorsed a "precautionary approach" and made some specific recommendations that it took to be in the spirit of the principle in the less stringent sense of taking steps to guard against unlikely or remote harm. Thus the refusal to license a GM maize variety carrying an antibiotic resistant marker gene as well as the general policy of conducting post-release monitoring could well be defended as justifiably prudent actions (see Nuffield, 1999, sections 2.48 and 7.40). All the same, the Working Party emphatically held that "concerns about very small risks

to the inhabitants of developed countries" did not prevail against what it saw as the "compelling" moral imperative for making GM crops readily and economically available to developing countries who want them.[3] This example shows that ethical considerations of a more specific character may enter into the balancing of risks and benefits and tilt the outcome in one direction or another.

THE DISTRIBUTION OF BURDENS OF PROOF

Advocates of free enterprise and anti-regulationists often consider the inauguration of the Precautionary Principle as an illicit attempt to introduce the maxim "guilty until proven innocent," that is, as an ominous and most arbitrary departure from the age-old legal principle "innocent until proven guilty." Write Miller and Conko: "In practice, the precautionary principle is interpreted to mean that a product or technology should be assumed to be guilty until its innocence can be proven to a standard demanded by its critics, leaving much arbitrariness and a standard that can seldom be met" (Miller and Conko, 2000).

Political scientist Wildavsky even discerns a monumental reversal of roles between state and citizen here: "The immensity of the change requires re-emphasis: private action requires proof of the absence of harm; governmental action requires no proof of harm. The relative role of the citizen and the state have been reversed. In the past it was the citizen who was entitled to act and the state that had to justify its intervention; now it is the state that intervenes by right and the citizen who has to give reasons for acting. The reversal of the usual course of action has profound implications" (Wildavsky, 1995, p. 430). We think that framing the discussion in terms of "the citizen versus the state" is rather rhetorical. The problem does not just concern the freedom of action of the individual citizen. The question is also whether private companies should be given carte blanche to undertake innovative activities that may possibly harm the welfare of (segments of) the general public. In a democracy, governmental regulation aims to protect the legitimate interests of the entire citizenry, and not simply to curtail individual freedom. Nevertheless, the grounds for intervention should be carefully defined.

[3] This position is criticized in Food Ethics Council, 1999. The contrast in positions is succinctly stated thus: "If Nuffield's position is described as 'Yes, but' (assuming there are no major problems, only minor ones), ours is 'No, unless ....' We believe there are significant drawbacks to GM food technology, which suggest it will only rarely be appropriate. Our response to GM foods is: 'No, thanks,' unless they serve vital roles with low risks."

Although it is true that the maxim "innocent until proven guilty" is a long-standing legal principle, there are also exceptions to this rule. On pragmatic grounds, certain areas of jurisprudence have introduced a so-called *reversal of the burden of proof* in specific circumstances. For example in German patent law, which does not protect chemical substances but only chemical processes, if company B markets a new compound that is already covered by a process patent granted to company A, the presumption is that company B has infringed the latter's patent. Company B can only free itself from this suspicion if it demonstrates before the court that it has found a new and independent chemical process enabling the production of the same compound, otherwise it will be found guilty of infringing company A's patent (appealing to a new process that is kept secret would be of no avail here). The reversal of the customary onus of proof was deemed necessary in cases such as this in order to ensure that the aims of patent protection for chemical inventions would not be completely defeated.[4] Another example is derived from Dutch civil law. A divorced wife who demands more alimony from her former spouse on the grounds that the latter has received a salary rise, does not always have to prove this fact. It is sometimes left to her former husband to disprove it. In view of banking secrecy, the burden of proof is thus placed with the person who has the lightest burden. Finally, the authors of the report *The politics of GM food: Risk, science & public trust* (a product of the ESRC Global Environmental Change Programme) provide an example derived from British law: "Under UK law, if your car is stolen and you claim that the thief damaged it, the thief must now prove otherwise to avoid liability. The act of the perpetrator makes it difficult for the victim to prove causation, so it is morally acceptable to switch the burden of proof" (ESRC Global Environment Change Programme, 1999). This example is taken by the authors of the report as an established legal precedent for a far-reaching proposal to revise the burden of proof in the area of environmental causation. We must now briefly examine this proposal.

The proposal aims to remedy the situation that the "perpetrators" of adverse impacts from new technologies often escape responsibility, because such impacts are hard to prove and the burden of proof is on those who suffer. The upshot of the proposal is explained thus: "Put simply, if an industry has released an environmental agent that is not safe, it is for the industry to prove that it was not the cause of any possible adverse impact" (ESRC, 1999, p. 19). Under the traditional burden of proof, those who suffer had to prove that an adverse impact occurred, that the environmental agent can cause the impact, and that the environmental agent was released.

[4] See about this problem in Zimmermann, 1965.

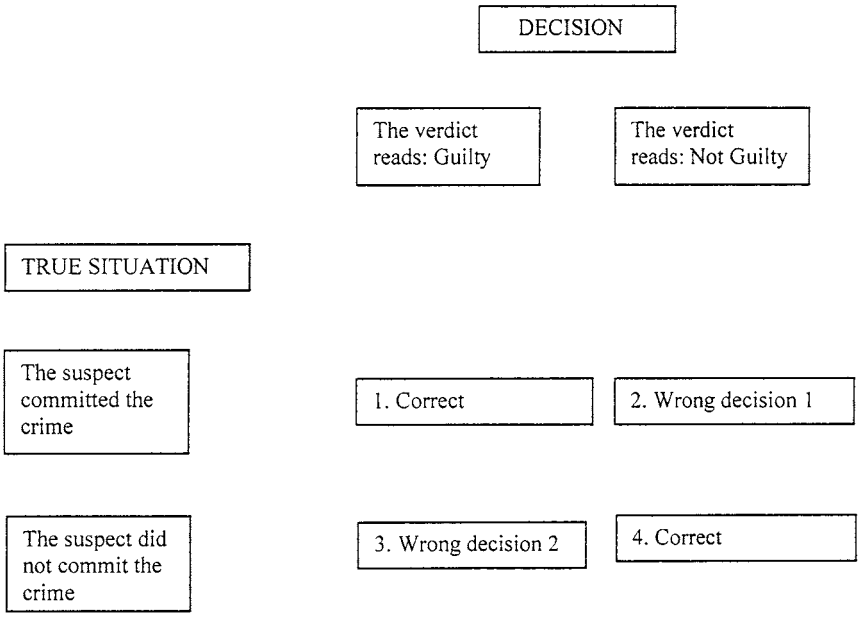
In addition the victim also had to prove that he was exposed to the agent, that the level of exposure could cause the injury, and that there was no other cause. Under the revised burden of proof that is proposed, those who suffer must still prove that an adverse impact occurred, that the environmental agent can cause the impact, and that the environmental agent was released. But here the burden of proof of the victim ends, for causation will now be deemed proven unless those who released the agent can show that the victim was not exposed to the agent, that the level of exposure could not cause the injury or that there was another wholly responsible cause. The authors of the report hold that implementation of this proposal in the area of environmental causation "would reduce the inclination for industries to 'chance it,' " and they add, "Responsible industries should welcome this change – it levels the playing ground between them and less scrupulous rivals." We believe that the prospective scope of operation of the revised burden of proof is still somewhat limited.

The formulation given suggests that the proposed revision of the onus of proof may be used in concrete liability suits, but only under the assumption that the potential harmfulness of the environmental agent in general is already well-established (otherwise the phrase "if an industry has released an environmental agent *that is not safe*" would beg the question!). As such, the proposal does nothing to resolve the latter, more difficult problem of establishing the safety or potential harmfulness of the agent. Nonetheless, it has been useful in stimulating our thought about the problem of the distribution of burdens of proof. What the concrete proposal shows is that talk about "the" burden of proof may be misleading. Contrary to the ingrained way of speaking, it is not logically necessary that we put the whole of this so-called "burden" on the shoulders of one party or the other; it turns out that we may divide this "burden" in parts and assign them to different parties. So we think we still have gained a welcome insight. Indeed, in this article we attempt to arrive at more refined ways of thinking about problems involving the distribution of burdens of proof.[5]

[5] In a comment on an earlier version of this article, one referee noticed that our exposition contains an oblique reference to liability and challenged us to respond to "the proposition that a well-crafted liability regime could do the work of the precautionary principle without being subject to the same kind of abuse (because it is an ex post rather than an ex ante remedy)." We think this is a very interesting challenge, but fear that a serious response would require the length of a full article. It is indeed true that product liability raises many of the same questions about the distribution of burdens of proof that are raised in the debate on the Precautionary Principle. We are reluctant, however, to devise a "well-crafted" liability regime. Standard texts in the economics of law emphasize that no one rule (e.g., strict liability, negligence, strict liability with a defence of contributory negligence, or some other variation) will be best in all the relevant respects like its effects on producer

This refinement can also be sought in a different direction. Traditional legal terminology suggests that the only matter at issue is where to place "the" burden of proof. As soon as we allow for more or less exacting *standards* of proof, however, an extra dimension of variation immediately becomes visible. In other words, the burden we want to put on the shoulders of one or the other party becomes more or less heavy depending on whether we set our standards of proof more or less highly. This consideration is very relevant to our present undertaking. It enables us to overcome and relativise the apparently clear-cut dichotomy between "innocent until proven guilty" and "guilty until proven innocent."

Let us re-examine the classical paradigm case of "innocent until proven guilty" in the context of criminal justice. Clearly, the opposite principle is rightly considered the epitome of arbitrariness and injustice and should have no place in a civilized democratic state in which "the rule of law" obtains. But matters may appear slightly more complicated if we look somewhat deeper into the basic problem any system of criminal justice has to confront. That problem can be visualized in the following matrix:



care, consumer care, risk allocation, and industry output (See, for instance, Polinsky, 1983). We also think it highly significant that in the shift towards a European strict liability regime, legislators have created special legal exemptions for so-called "development risks" (for new products) in order not to stifle innovation. We have seen a similar concern among some of the critics of the Precautionary Principle. For more on product liability, see Leen, 1999.

The matrix shows that there are two main ways in which a miscarriage of justice can come about. Either the suspect did not commit the crime, but the verdict found him guilty; or the suspect did commit the crime, but the verdict found him not guilty. In a civilized system of justice, the risks of the first type of error are minimized as far as possible. Such a system contains safeguards and precautions in the form of high standards of proof so as to ensure that a suspect will be condemned for a certain criminal offence only if it has been established "beyond reasonable doubt" that he in fact committed the alleged offence. Alas, there is a price to be paid for this cautious and civilized approach, namely the possibly large number of wrongdoers who have to be acquitted due to "lack of sufficient proof." To a certain extent, the risks of the two types of error are inversely related. We may try to reduce the risk of condemning an innocent person by demanding ever more exacting standards of proof, but only at the expense of increasing the risk of acquitting culpable offenders. So we must recognize that there is an inevitable *trade-off* involved in the design of our system of criminal justice. We may attempt to set our standards as high as we can, but somewhere a balance must be struck, lest the system will become unworkable by making it too difficult to pass sentence on the majority of wrongful offenders.

Considerations such as these are also helpful, we think, to clarify the confusing debates engendered by the Precautionary Principle.

"SOUND SCIENCE" ONLY?

According to Henry I. Miller and other advocates of minimal regulation, the only grounds providing a sufficient case for regulating new technologies are valid reasons based on "sound science." In their view, considerations derived from some interpretation of the Precautionary Principle or from worst-case analyses are inherently biased, unscientific, and therefore inadmissible as grounds for regulation. The British investigators of the ESRC Global Environmental Change Programme take precisely the opposite view. They hold that due to the inevitable selectivity in framing assumptions, science is inherently incapable of dealing adequately with the profound uncertainties that are raised by biotechnology and other new technologies: "We [. . .] suggest that science cannot provide definitive answers in these cases, so the policy of relying on claims of 'sound science' may, ironically, itself be unsound. Ethical issues are central" (ESRC, 1999, p. 4).

It is not very likely, however, that countries will openly confess that their regulations are based on more than just "sound science," for they

would thereby expose themselves to the charge of instituting protectionist measures and thus to the risk of retaliatory actions from other countries. International agreements on setting product standards and resolving trade disputes still rely very much on the idea that valid scientific reasons alone may provide legitimate grounds for regulation. The UN Codex Alimentarius Commission, for example, which sets international food standards and is the final arbiter where trade disputes depend on health issues, is only permitted to consider "valid" scientific reasons before products can be prevented from entering the market in a member country. But what is a "valid scientific reason?" It is interesting to note that precisely this question was also raised by the British House of Commons Select Committee on Science and Technology in its First Report on the Scientific Advisory System: "[this] raises the questions of what is a valid reason, and who should judge?" Furthermore, the Select Committee made a specific recommendation on this point: "We recommend that the Government seeks to establish international agreement on what constitutes a 'valid' scientific reason; *and that the definition of validity is based on the precautionary principle* [our emphasis]" (House of Commons, 1999). Understandably, such efforts are regarded with the utmost suspicion by the advocates of minimal regulation. Miller and Conko charge that "the EU and environmental activists are trying to undermine the WTO [...] by writing the precautionary principle into the standards of the Codex Alimentarius Commission" (Miller and Conko, 2000, p. 12), which in their view would constitute a deplorable departure from "science-based regulation."

Thus the scene before us is one where various parties are busily engaged in all kinds of "boundary work":[6] where to draw the line between "science" and "politics," or between "risk assessment" and "risk management"? And what sorts of "precautionary" considerations may legitimately enter into any of these different undertakings?

One political body struggling with these issues is the European Commission. In its position paper the Commission introduced a distinction between a "prudential approach [...] which is [...] an integral part of the scientific opinion delivered by the risk evaluators," on the one hand, and "application of the precautionary principle [which] is part of risk management," on the other (European Commission, 2000, p. 13). Risk management is the preserve of political decision-makers, according to the Commission. It takes what seems to be a firm stand on this issue: "Judging what is an 'acceptable' level of risk for society is an eminently *political* responsibility" (European Commission, 2000, p. 4).

[6] For the notion of "boundary work," see Gieryn, 1995.

It is clear that the Commission intends to draw a sharp boundary between science and politics, or between "risk assessment" and "risk management." Still, we may wonder whether various "transscientific" (hence "political") considerations already necessarily enter into the "prudential approach" used in scientific risk assessment. According to the explanations given in Appendix III of the Commission's report, "worst-case" assumptions may be admissible in risk assessment when the available data are inadequate or insufficient. But can the choice of such assumptions be justified solely on scientific grounds? Doesn't the scientific risk analyst thereby anticipate and pre-empt the decision that was understood to be the preserve of the politician, namely to decide on whether a given risk is "acceptable" to society?

In the following passage, Wildavsky refers approvingly to the efforts of some unnamed "analytical purists" who want to maintain a strict distinction between "risk assessment" and "risk management."

> There are analysts who want to rescue risk analysis from the multiplication of worst cases by arguing for analytic purity. By assuming the worst, with uneven definitions of how bad that can be, these purists claim, risk assessment becomes confused with risk management. They recommend that the least-biased risk assessment be performed and then the degree of caution desired be applied by administrative managers or elected political officials. I agree that this separation of science from politics, though it cannot be entirely achieved, is worth striving for. After all, if it is all politics, why bring in science?' (Wildavsky, 1995, p. 431).

It may indeed be desirable to separate science from politics, but the proposed distinction is more wishful thinking than reality, as Wildavsky himself also seems to realize. After all, what would be *the least-biased* risk assessment? Is the worst case assumption always the most biased? Is an assumed somewhat-less-than-worst case always less "biased" than the worst case, and a much-less-than-worst case even more less biased? Who should decide, and on what grounds?

The Dutch philosopher Ad van Dommelen claims to be able to "distinguish a *political* burden of proof for making claims on 'acceptability' from a *scientific* burden of proof for making claims on 'empirical adequacy' or 'plausibility' " (Dommelen, 1999, p. 91). He disagrees with the German philosopher Wolfgang van den Daele, who holds that it is a genuinely political issue "on which side the burden of proof for the safety of a technology should be put." Like his political antipodes Miller and Conko, who swear by "sound science," van Dommelen attempts to find solutions for problems of hazard identification and risk assessment exclusively within the realm of science. He argues that the applied science of biosafety assessment should take the utmost care to define an adequate *problem definition*, or *set of relevant research questions* (SRQ), which constitutes a so-called *window of concern*. Although he asserts that both those who

want to argue for the *inclusion* of a certain research question within the problem definition and those who want to argue for its *exclusion* thereby assume a (scientific) burden of proof, his SRQ-methodology has clearly loaded the dice in favor of the first group: "In practice, the burden of proof for this type of applied science will be on inclusion as well as on exclusion of possibly relevant research questions for identifying GEO [GMO] hazards. The purpose of the research (identifying hazards and assessing biosafety) implies that in cases of reasonable doubt, a possibly relevant question must be *included* in the applied SRQ. Any other methodology of dividing the burden of proof for claims on GEO hazard identification would impair the scientific legitimation of biosafety claims" (Dommelen, 1999, p. 65). Because in this way it becomes fairly easy to argue for the inclusion of some allegedly "relevant" research question, and extremely difficult if not impossible for its exclusion, the *window of concern* has a tendency to expand without limit. It is no coincidence that in all the several controversies related to GMO biosafety assessment that he examines in his book, van Dommelen consistently comes down on the side of those who are critical of biotechnology and concerned about its risks. The main defect of his SRQ-methodology is that it hardly provides for a possibility to narrow down the "window of concern." He defines the taking up of the *political* burden of proof rather tendentiously as the decision to "overrule" the scientific viewpoint.[7] It seems that those who argue for the extension of the "window of concern" always have scientific reason on their side; those who argue for narrowing this window must act from "unscientific," political motives.

Sometimes GMO biosafety assessment will involve statistical tests.[8] Van Dommelen holds that also the implementation and interpretation of such tests can proceed entirely within the realm of science, without invoking any "transscientific" considerations. We reject this conclusion

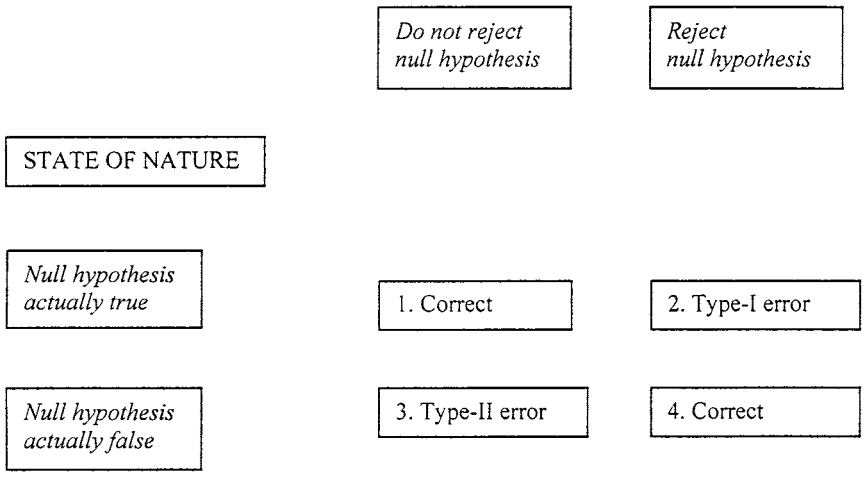
[7] Dommelen, 1999, p. 92. On p. 151 van Dommelen criticizes representatives of the COGEM, the competent authority on GMO biosafety assessment in the Netherlands, for not accepting what he sees as the full set of relevant questions: "Where the COGEM representatives decide not to take on the *scientific* burden of proof for excluding the possibly relevant research questions as suggested by Middelhoven (1997) from a sufficient SRQ for the specific research purpose, they thereby take on the *political* burden of proof for arguing the acceptability of using a less than scientific basis of legitimation for their review of notifications."

[8] One referee commented on an earlier version of this article that "statistical tests are not the only way in which science can make a contribution to risk assessment." We agree. The reason for entering into an analysis of statistical testing is that we can demonstrate most clearly in this purportedly purely technical field the salience of a particular distribution of the burdens of proof, as embodied in conventional significance levels reflecting the relative importance attached to errors of type I and type II.

as mistaken and believe that it is based on misunderstanding the logic of statistical tests. Even more, we think that conducting a statistical test always involves a particular (albeit often only implicit) distribution of the burdens of proof. We therefore hold that a more thorough examination of the structure of statistical inference will vindicate our general position. So let us have a closer look.

In the practice of statistical testing, it is customary to formulate a so-called *null-hypothesis*, H_0, as an alternative to the proper *research hypothesis*, H_1. Because the null-hypothesis is usually stated in terms of "no effect," it is much easier to put to the test. If the outcome of the test warrants the rejection of the null-hypothesis, the researcher will normally accept the alternative hypothesis H_1. As always in statistical reasoning, both the decision to reject and to accept a particular hypothesis is associated with a certain risk.

The following matrix (which not entirely by chance resembles the previous matrix about the basic problem of criminal justice) shows the four possibilities with testing the null-hypothesis.



In the final chapter of his book, van Dommelen also enters into a discussion of statistical testing. He notes that in statistical tests aimed at the assessment of ecological risks (e.g., the inadvertent spread of transgenic organisms into the environment) type-I errors usually get the most attention, but that type-II errors (in this case: failures to reject the false null hypothesis of no spread of the transgenic organism) "are no less interesting in the context of a precautionary hazard identification" (Dommelen, 1999, p. 187). He also notes, following Peterman and M'Gonikle, that statements

about safety (or lack of an important effect) only make sense when accompanied with a specification of the risk of making a type-II error, but that most empirical studies fail to mention the associated level. These points are well taken. Deeply problematic, however, is van Dommelen's subsequent proposal to deploy his general SRQ-methodology in order to get rid of type-II errors altogether: "[. . .] Type-II errors of statistical interpretation must be recognized and avoided by giving due attention to the relevant empirical questions for a specific research purpose" (Dommelen, 1999, p. 187). Actually, his SRQ-methodology has not much to do with the recognition of type-II errors, and even less with their avoidance. Indeed, it would be absurd to demand that such errors be *avoided*! The whole point of statistical testing is that the possibility of such errors can never be avoided. At best, the risk of such errors can be reduced (to a certain extent and subject to constraints) and, ideally, we can calculate the level of the risk involved. Statistics is the art of taking "calculated" risks.

Van Dommelen concludes the section on statistical testing with the following paragraph: "Possible *methodological pitfalls* such as these [i.e., ignoring type-II errors] can only be recognized and avoided by the scientific experts who are involved with practical problems such as biosafety assessment. Only scientists can provide the required safeguarding from these pitfalls. Without their qualified input, those responsible run the risk of dealing with *scientific* questions as if they were *transscientific*" (Dommelen, 1999, p. 188). A better understanding of statistical reasoning would have allowed him to draw precisely the opposite conclusion. Every elementary handbook of statistics will tell you that there is a trade-off between type-I and type-II errors.[9] You may attempt to reduce the one but at the cost of increasing the other. Usually the probability of making a type-I error – the so-called *significance level* – is set at a conventional level, say at the 0.05, 0.01, or 0.001 level ("there is nothing sacred or absolute about these levels!" as our statistics handbook emphasizes), and, given the use of a particular test, this decision brings with it a specific probability of making a type-II error (which, however, is often more difficult to calculate). It is worth repeating that these levels are conventional through-and-through. If, because of the environmental interests at stake, we would be extremely concerned about the possibility of failing to reject the null hypothesis of "no effect" when it is actually false, we could indeed devise our statistical tests so as to minimize the risk of this error. But only at a cost, namely that the probability of making type-I errors will be increased. If we set out to miss hardly any possible effect, we will inevitably make a lot of type-I errors, which means in this context that we will produce a lot

[9] See for instance Blalock, 1960, pp. 122–128.

of false ecological alarms. Van Dommelen likes to see science as our major "danger detective" but is rather blind to the flip side of this undertaking.

The important point to note in this connection, however, is that before we can design appropriate statistical tests, we have to make up our minds as to the (economic and moral) costs associated with either of the two types of error. If we are particularly concerned about the possible spread of transgenics into the environment and attach much importance to avoiding such an occurrence, we will consider "false negatives" extremely costly. Our decision to substantially reduce the probability of making type-II errors will then automatically increase the burden of proof of those who want to show that the risks of a certain biotechnological application are acceptable. Conversely, if we are eager not to miss out on any of the attractive economic (but perhaps also ecological) opportunities opened up by new biotechnological applications, we may want to reduce the chance of incurring type-I errors accordingly, thus imposing a so much larger burden of proof on those who want to argue that these opportunities must be foregone because of the serious ecological hazards involved. Even such widely employed routines in scientific research as the use of standard significance levels in statistical testing, far from being a purely technical matter, implicitly embody a particular distribution of the burdens of proof. How we divide these burdens is a genuinely political issue. Thus the "politics" of risk management has already penetrated the core of the allegedly purely "scientific" risk assessment. Simply to call on "sound science" to solve our problems will therefore not do.

If *science* cannot possibly provide an answer to our problem of how to divide the burdens of proof, perhaps we should finally turn to *ethics*. Kristin Shrader-Frechette believes that a general ethical case can be made for preferring type-I errors to type-II errors in situations with potentially serious consequences (Shrader-Frechette, 1994, pp. 101–117). She holds that in such situations "the burden of proof (regarding risk acceptability) should be placed on the person wishing to reduce [type-I, rather than type-II, risk]" (Shrader-Frechette, 1994, p. 111), that is, in our case, on the person wishing to implement a certain technological innovation. In her view, protecting the public from serious harm (e.g., loss of species, nuclear accidents) takes precedence over enhancing welfare (e.g., by developing new agricultural crops or by providing electrical power on demand). We do not find her moral case entirely convincing, however. For one thing, the suggested dichotomy between protecting from harm and enhancing welfare is far from clear-cut, as she herself also admits.[10] Secondly, the

[10] Shrader-Frechette, 1994, p. 112: "Of course, it is difficult to draw the line between what enhances welfare and what avoids harms."

issue is not simply whether type-I errors or type-II errors have to be *minimized* (taken in an almost absolute sense), but rather where to strike a reasonable balance. One can of course legitimately argue that type-II errors should be given more weight and attention than is usually accorded them, especially in situations with potentially serious consequences, but the logic of statistical inference does not support an absolute precedence of one type of error above the other. After all, it stresses that there is always a *trade-off* between the two error types. Therefore, and in line with our general argument, the issue is not whether the burden of proof should be placed on the developers of new technologies or on their opponents, but how it should be divided among the two parties.

CONCLUSION

In this article, we suggested reframing the debate on the significance and the scope of the Precautionary Principle into a debate on the proper division of burdens of proof. In the existing debate there is a polarization between "innocent until proven guilty," on the one hand, and "guilty until proven innocent," on the other. Under the traditional burden of proof, those who suffer from an innovation have to prove that an adverse impact occurred. This may lead to irreparable damage to society. Adherents of the Precautionary Principle defend a reversal of the burden of proof. This may lead to an early halt to innovation.

We attempted to arrive at a more refined way of thinking by replacing the dichotomy between the two kinds of burden of proof by a system that distributes different kinds of burdens of proof. We elaborated this system in three different directions:

1. a philosophical distinction between misfortunes of the first kind and misfortunes of the second kind;
2. a legal distinction between two kinds of wrong decisions;
3. a statistical distinction between type I-errors and type II-errors.

In all these directions, it was found that we have to look for a balance in the system of burdens of proof because we must recognize that there is an inevitable trade-off involved. The two kinds of error are a system of communicating vessels: diminishing one kind of error leads to increase of the other.

Finally, we discussed a few possibilities for future research. Firstly, while the traditional burden of proof and the burden of proof under the Precautionary Principle rely on a distinction between science and politics, the recognition of an inevitable trade-off leads to further research into their

inter-relatedness, and it is no longer possible to draw a sharp boundary between risk-assessment and risk-management. Secondly, it is possible to decide on political or ethical grounds to give priority to avoiding one type of error, but this priority can never be absolute.

REFERENCES

Blalock, H. M., *Social Statistics* (McGraw-Hill, London, 1960).

Conko, G., "Throwing Precaution to the Wind: The Perils of the Precautionary Principle," CEI Update, August/September (available on the website of the Competitive Enterprise Institute: http://www.cei.org), 2000.

Dommelen, A. van, *Hazard Identification of Agricultural Biotechnology: Finding Relevant Questions* (International Books, Utrecht, 1999).

ESRC Global Environmental Change Programme, *The Politics of GM Food: Risk, Science & Public Trust*, Special Briefing No 5, October (available on the Internet: http://www.susx.ac.uk/Units/gec/gecko/gm-brief.htm), 1999.

European Commission, *On the Precautionary Principle*, COM 1, 2000.

Food Ethics Council, *Novel Foods: Beyond Nuffield* (accessible on the website of the National Council for Biotechnology Education: www.ncbe.reading.ac.uk/ncbe/gmfood), 1999.

Gieryn, T. F., "Boundaries of Science," in S. Jasanoff et al. (eds.), *Handbook of Science and Technology Studies* (Sage Publications, Thousand Oaks/London/New Delhi, 1995).

House of Commons, Select Committee on Science and Technology, *First Report* (available on the Internet: http://www.parliament.the-stationary-office.co.uk/pa/cm199899/cmselect/cmsc. . ./286 06.ht), 1999.

Kellow, A., "Risk Assessment and Decision-Making for Genetically Modified Foods," *IPA Biotechnology Backgrounder*, Number 1, October 1999.

Leen, A., *The Consumer in Austrian Economics and the Austrian Perspective on Consumer Policy*, Wageningen (thesis), 1999.

Losey, J. E., L. S. Rayor, and M. E. Carter, "Transgenic Pollen Harms Monarch Butterfly," *Nature* 399 (1999), 214.

Miller, H. I. and G. Conko, "The Science of Biotechnology Meets the Politics of Global Regulation," *Issues in Science and Technology On Line*, Fall (available on the Internet: http://www.nap.edu/issues/17.1/miller.htm), 2000.

Nuffield Council on Bioethics, *Genetically Modified Crops: The Ethical and Social Issues* (available on the Internet: http://www.nuffield.org/bioethics/publication/modifiedcrops), 1999.

Polinsky, A. M., *An Introduction to Law and Economics* (Little, Brown and Company, Boston and Toronto, 1983).

Rescher, N., *Pluralism: Against the Demand for Consensus* (Clarendon Press, Oxford, 1995).

Schmickle, S., "Researchers Say Bt Corn Poses Low Risk to Monarch Butterflies," *Star Tribune*, 18 November, 2000.

Shrader-Frechette, K., *Ethics of Scientific Research* (Rowman & Littlefield, Lanham and London, 1994).

Wildavsky, A., *But Is It True? A Citizen's Guide to Environmental Health and Safety Issues* (Harvard University Press, Cambridge, MA, 1995).
Zimmermann, P. A., *Patentwesen in der Chemie. Ursprünge, Anfänge, Entwicklung* (Ludwigshafen, 1965).

Applied Philosophy
Hollandseweg 1
6706 KN Wageningen
E-mail: Henk.vandenBelt@alg.tf.wag-ur.nl
Bart.Gremmen@alg.tf.wag-ur.nl