

Spring 2003

The Myth of Science as a 'Neutral Arbiter' for Triggering Precautions

Vern R. Walker

Maurice A. Deane School of Law at Hofstra University

Follow this and additional works at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship

Recommended Citation

Vern R. Walker, *The Myth of Science as a 'Neutral Arbiter' for Triggering Precautions*, 26 B.C. Int'l & Comp. L. Rev. 197 (2003)
Available at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship/150

This Article is brought to you for free and open access by Scholarly Commons at Hofstra Law. It has been accepted for inclusion in Hofstra Law Faculty Scholarship by an authorized administrator of Scholarly Commons at Hofstra Law. For more information, please contact lawcls@hofstra.edu.

THE MYTH OF SCIENCE AS A "NEUTRAL ARBITER" FOR TRIGGERING PRECAUTIONS

VERN R. WALKER*

Abstract: This article demonstrates that science cannot be a "neutral arbiter" for triggering precautionary measures, because both making and warranting findings of risk require non-scientific decisions. Making a risk finding requires decisions about the meaning of "risk of harm," about the meaning of any modifiers for that predicate, and about the degree of confidence asserted for the finding as a whole. Determining that the available scientific evidence warrants a finding of risk requires decisions about acceptable degrees of various types of uncertainty—namely, conceptual uncertainty, measurement uncertainty, sampling uncertainty, modeling uncertainty, and causal uncertainty. This article illustrates these decisions using examples from the food safety law of the United States, recent animal feed cases in the European Community, and Appellate Body decisions in WTO trade disputes. Finding a risk that triggers precautions cannot be a purely scientific act, notwithstanding the myth that a "value-neutral" science can do so.

INTRODUCTION

It is tempting to think that scientists, acting purely as scientists, can make the risk determinations that would trigger the taking of precautions. If this were true, then perhaps decision-makers could remove value-laden politics from the factfinding processes that ground the domestic regulation of health and the environment,¹ the invocation of the precautionary principle under treaties,² and the adjudication of international trade disputes.³ Decision-makers could

* © 2003, Vern R. Walker, Professor of Law, Hofstra University; Ph.D., University of Notre Dame, 1975; J.D., Yale University, 1980. The author is grateful for the financial support provided by a research grant from Hofstra University.

¹ For examples of risk findings concerning food safety, see Section III.A *infra*.

² For an example involving the precautionary principle of European law, see Section III.B *infra*. For a general discussion of triggering precautionary measures under various treaties, see generally James E. Hickey, Jr. & Vern R. Walker, *Refining the Precautionary Principle in International Environmental Law*, 14 VA. ENVTL. L. J. 423 (1995).

³ See, e.g., discussion *infra* Section III.C.

make a clean division between determining whether a risk exists at all (risk assessment by scientists), and deciding what actions to take about such risks (risk management by governmental decision-makers). The thesis of this article, however, is that making such a risk determination cannot be a matter of "pure science." Although science can and should play an essential role in guiding the warranted findings about risk that trigger the taking of precautions, such factfinding necessarily includes decision-making that cannot be purely scientific in nature. At least in the most important cases, it is a myth that science can be a "neutral arbiter" for triggering precautionary measures. My argument has two dimensions. The first concerns the logical structure of the risk findings. The triggering finding requires decisions that cannot be "purely scientific." This means that when lawmakers *define* the factual predicate for taking precautions, non-scientific decisions are necessary. It also means that when factfinders *identify* risks, or *make* a finding that the legal trigger for taking precautions has been satisfied, non-scientific decisions are also necessarily involved. In theory, it is conceivable that when lawmakers establish the wording of the relevant finding, they could use terms that have no uncertainty as to their meaning, and factfinding using those words could be a ministerial task. In practice, however, quite the opposite is true: lawmakers often use vague terms, and finding that the legal trigger about risk has been met requires decision-making that cannot be purely scientific.⁴ Unfortunately, lawmakers and factfinders often mask non-scientific decisions in language that sounds scientific.

The second dimension of my argument concerns the logical relationship between available scientific evidence and findings of risk. Inherent in the evidentiary warrant for the finding are scientific uncertainties, which require decisions about acceptability that cannot be "purely scientific." Deciding that a particular finding about the risk is warranted, given the evidence, cannot be a policy-neutral determination. Once again, factfinders, and even scientists, sometimes mask those non-scientific decisions about warrant in language that sounds scientific.

It is a myth, therefore, to think that science can be a "neutral arbiter" for triggering precautions. Moreover, risk assessment, defined as the process of reaching warranted findings about risk on the basis of scientific evidence, cannot be purely scientific or policy-neutral. This article develops this thesis in three parts. The first section of the

⁴ For examples of vague formulations in use today, see *infra* Section III.

article discusses the logical structure of a finding about risk, and analyzes the logical elements of such a finding. The second section surveys the types of scientific uncertainty that are logically inherent in any inferences from empirical evidence to risk findings. The third section illustrates the way in which these logical elements and scientific uncertainties underlie legal language that often sounds scientific. The illustrations come from the food safety law of the United States, the animal feed law of the European Community, and the health and safety provisions of a treaty administered by the World Trade Organization (WTO). The variety displayed in these disparate legal areas emphasizes the logical basis of my thesis. Because my analysis is grounded in logic—that is, in the meaning of the risk finding itself and in the nature of its scientific warrant—the conclusion is inescapable that warranted risk findings used to justify the taking of precautions cannot be purely scientific.

I. NON-SCIENTIFIC DECISIONS INHERENT IN THE LOGICAL STRUCTURE OF FINDINGS ABOUT RISK

The first dimension of my thesis is the logical structure of the findings about risk. When lawmakers establish the wording for such a finding, they encounter a number of decisions that are not “purely scientific.” Moreover, when lawmakers use words that are necessarily or conveniently vague in their meaning, factfinders must make non-scientific decisions when they find that the legally required trigger for taking precautions has been met. The semantic structure of such findings about risk reflects those decisions. A finding of risk is some variant of the proposition “there exists a risk of harm.” It is an assertion that has three logical elements: (1) a categorical predicate (expressed in this example by the noun phrase “risk of harm”); (2) a modifier of the predicate (in this example, the indefinite article “a”); and (3) the modality or degree of confidence in the truth of the proposition as a whole (here, expressed in part by the indicative verb “exists”). I will briefly discuss each logical element in turn.

First, the finding about risk must employ some categorical predicate, which the factfinder will use to classify any particular situation under consideration. Precautions are taken, in the face of risk, in an effort to prevent some harm from occurring. In this general analysis, “harm” can mean any adverse event. What is considered as “adverse” usually involves an evaluation that is not entirely scientific. Scientists might be able to determine which conditions are harmful to a particular organism or ecosystem if “harm” means clear biological dys-

function. The matter becomes more complicated, however, if what counts as "harm" or "adverse" requires evaluation and balancing of psychological, economic, or aesthetic effects. If "harm" connotes a net loss when weighed against benefits, there may be no consensus on what to value or count as a benefit, or on acceptable trade-offs. In such a case, "harm" is no longer a term with a scientific definition. Moreover, if "harm" is a matter of degree, then deciding what constitutes a legally recognized "harm" requires establishing a threshold on a scale of degree—for example, a threshold of air or water quality short of clear toxicity. A definition of the *kind* and *degree* of "harm" that *should* trigger precautionary measures is necessarily a balancing of risks and benefits under the circumstances, which is beyond the domain of pure science. Thus, if lawmakers simply employ a vague term such as "harm" in the legal definition of the required finding about risk, then they have passed on to factfinders the non-scientific task of deciding what to count as "harm" for legal purposes.

A similar analysis applies to the meaning of "risk" in the categorical predicate, which can connote either the possibility of a harmful event or some measure (probability) of the likelihood that the harm will occur.⁵ While the "harm" refers to the adverse end-state, the concept of "risk" connotes a causal chain of events leading to that end-state. A chain of events denoted as "a risk" could begin with the speaker's situation ("I am at risk for lung cancer") or action ("by smoking cigarettes I am taking a risk of developing lung cancer"), or could begin with any other situation or action ("releasing this genetically modified organism into the environment is creating a risk of ecological harm"). Such event-chains can be either direct or indirect (mediated by other events), and intervening events can be known or unknown. The causal links between events might be described either qualitatively or quantitatively. If a causal link is described qualitatively, then the speaker is asserting only that the causal chain of events is a possible one. For example, "cigarette smoking increases the risk of lung cancer" asserts that there exists at least one causal chain that leads from smoking cigarettes to developing lung cancer, a causal chain that is not a possibility in the absence of cigarette smoking. The intervening causal mechanisms may not be known, but there can be

⁵ See, e.g., THE RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE 1660 (2d ed. 1987) (defining "risk" as "exposure to the chance of injury or loss"); WEBSTER'S NINTH NEW COLLEGIATE DICTIONARY 1018 (1991) (defining "risk" as "possibility of loss or injury"); OXFORD ENGLISH DICTIONARY ONLINE (defining "risk" as "exposure to mischance or peril"), at <http://www.oed.com> (last visited Apr. 8, 2003).

good evidence of the causal link nonetheless. A causal link might also be described quantitatively. For example, the speaker might assign a conditional probability to the likelihood of occurrence—such as a 0.3 probability of the harm occurring (“H”) if the action (“A”) is taken (in symbols, $\text{Prob}(H|A) = 0.3$). These are all logical options for how lawmakers might define a triggering “risk.”

Finding a qualifying “risk of harm,” therefore, combines a decision about what constitutes a “harm” with an assertion about the possibility or likelihood of causal chains resulting in that harm. Science certainly has a critical role to play in identifying and, if possible, quantifying those causal chains. But in the legal context of triggering precautions against harm, the word “risk” generally means a particular *level* or *degree* of risk that *should* trigger the taking of precautionary measures. Selecting the triggering level of risk requires a balancing of values, *i.e.*, a weighing of the costs of precaution against the costs of risk-taking. This is clearly the case when risk is measured quantitatively, and there is some notion of *de minimis* risk—that is, some probability of occurrence that is so low that it is “not considered a risk at all” in the relevant legal context. When lawmakers establish that a threshold of risk must be cleared before precautions are triggered, this decision is not scientific. When lawmakers simply employ the word “risk” to formulate the legally required trigger, then they have passed on to factfinders the non-scientific task of deciding what to count as a qualifying “risk.”

Second, a statement or finding of trigger for taking precautions also may involve, implicitly or explicitly, a modifier for the categorical predicate. Sometimes the selection of a modifier simply reinforces the notion that there is a risk threshold below which the precautions are not triggered. For example, the statement of trigger might specify that only a “serious” or “substantial” risk of harm should trigger the precautions. In other cases, the statement of trigger might refer to “a” risk, or even “any” risk. In such situations, a fair reading of the requirement might be that any possible chain of events leading to a relevant harm is sufficient to trigger precautions. Thus, lawmakers might insert either a qualitative or quantitative modifier for the categorical predicate. In the absence of linguistic signals prohibiting the practice, factfinders might infer an implicit modifier. For example, they might decide that a risk of harm is so remote or so trivial that it is not a “real” risk at all—at least for purposes of taking the precautions at issue in the legal context. Regardless of which modifier lawmakers select for the legal standard, the selection itself clearly is not a scientific decision. Moreover, if the selected modifier is vague, or calls

for some balancing of multiple factors, then deciding whether it is met in a particular case cannot be a purely scientific decision. For example, it is not a scientific matter to decide whether a risk is "serious" or "substantial" *enough* to trigger precautionary action.

Third, any finding explicitly or implicitly involves a degree of confidence in how likely it is to be true. This might be expressed simply in the grammatical mood of the verb used. The indicative and subjunctive moods in English express many variations. For example, a finding can assert that a risk of harm "does exist," or "might exist," "would exist," or "could exist." Logicians sometimes describe modalities as ranging from the (necessarily) false, through degrees of improbability, past a region of equipoise (as likely as not), through the increasingly probable, to the (necessarily) true.⁶ Examples of English expressions marking points along that range include "merely possible," "somewhat likely," and "highly probable." Lawmakers can decide to require any particular modality or degree of confidence before precautions are triggered. Moreover, they can match any particular risk of harm with any degree of confidence. For example, they might require precautions if a factfinder finds a "substantial risk of cancer" to be merely "possible," indicating some low level of confidence in the finding. By contrast, lawmakers might intend to trigger precautions if "some risk of any harm" has a "high likelihood of being true," indicating that the factfinder should be quite sure that some risk exists, although that risk might be fairly minor in nature.

The selection of which modality to include in the trigger clearly is not a scientific matter. The degree of confidence that is appropriate for triggering precautions depends upon a prudential balancing of the risks and benefits in the situation. It is not an issue for scientists alone to determine. Nor is it a purely scientific matter to determine whether the selected modality is satisfied in the particular case, as long as the relevant modality is not a purely scientific concept. For example, if the statement of trigger contains the modality "highly likely," then deciding what counts as "highly likely" is not a purely scientific matter. What should count as a "high" likelihood for purposes of risk taking is not a purely scientific question, and vague, imprecise terms do not become scientific simply because scientists are willing to

⁶ For discussions of modal logic, see, e.g., CLARENCE I. LEWIS & COOPER H. LANGFORD, *SYMBOLIC LOGIC* 153-66, 199-234 (2d ed. 1959); GERALD J. MASSEY, *UNDERSTANDING SYMBOLIC LOGIC* 163-222 (1970); MARK SAINSBURY, *LOGICAL FORMS: AN INTRODUCTION TO PHILOSOPHICAL LOGIC* 13-19, 220-90 (1991).

use them. Even if the evidence warranted assigning mathematical probabilities to various risks, there still would be the question of what should count as “highly likely”—probabilities of at least 0.7, or 0.85, or only 0.9? Perhaps this is better illustrated on the low end of the probability spectrum. If precautions are not triggered unless risks are “more than merely speculative,” surely it is not a scientific task to decide the meaning of this standard, or the degree of confidence needed to satisfy it. Statutes and regulations hardly ever define degrees of confidence in such a way that they become precise, scientific concepts.

This analysis applies to any triggering finding of risk. No matter what the specific statement of trigger is, it must implicitly or explicitly involve decisions about the meaning of the categorical predicate (such as “risk of harm”), about the meaning of any modifiers for the categorical predicate, and about the modality of the triggering proposition as a whole. When lawmakers make these decisions in formulating the legal statement of the triggering risk, they balance principles, policies, and the desirability of possible consequences. When enacted statements of trigger include words that are vague in meaning, then finding that any particular situation satisfies that trigger cannot be a matter for pure science. Such statements of triggering conditions simply push the non-scientific decision-making onto the factfinder, in the context of the particular case. Section III of this article contains examples of such vague statements of triggering risk.

II. NON-SCIENTIFIC DECISIONS INHERENT IN THE SCIENTIFIC WARRANT FOR FINDINGS ABOUT RISK

The previous section analyzed the logical structure of any statement that a triggering risk exists—either the kind of statement that a lawmaker might use to legislate the conditions for triggering precautionary measures, or the kind of statement that a factfinder might assert in finding those conditions to be satisfied in a particular case. In this section, I argue that when a factfinder concludes that the evidence warrants a finding of risk, this act necessarily involves non-scientific decision-making. This argument is logical in nature because it rests on showing that even scientific evidence about causation, of the kind used to warrant a finding of risk, necessarily involves several distinct types of uncertainty. Scientists sometimes can reduce the levels of such uncertainties, but they can never eliminate those uncertainties altogether.

In reaching a conclusion that empirical evidence warrants the proposition that "events of kind A *can cause* events of kind B," a reasonable person must decide, with respect to various types of uncertainty, what level of uncertainty is acceptable. Decisions about the acceptability of different levels of uncertainty—while they are often made by scientists, either ad hoc or by convention—are not themselves purely scientific decisions. They are pragmatic in nature and usually context-dependent. This section therefore provides the second dimension of my argument: any determination that the available scientific evidence *warrants* the triggering finding of risk necessarily involves non-scientific decision-making. In practice, these decisions may take the form of rules or findings about the acceptable methodological quality of the supporting evidence, the required specificity of the evidence, or the minimum sufficiency of the total amount of evidence. Such rules or findings reflect non-scientific decisions about when it is desirable to take precautions in the face of uncertainty about causation.

A. *Five Types of Scientific Uncertainty in Warranting Risk Determinations*

Empirical scientific evidence warranting any finding of risk involves at least five logically distinct types of scientific uncertainty: conceptual uncertainty, measurement uncertainty, sampling uncertainty, modeling uncertainty, and causal uncertainty.⁷ "Uncertainty" here means a potential for error in drawing an inference. Each type of uncertainty arises at a distinct step in the scientific method for warranting a finding about causation. After the process of scientific proof is complete, a residual degree of uncertainty of each type is inherent in the warrant for the finding. The argument here is that these types are logically distinct, generally cumulative, and probably inherent in every important finding of a triggering risk. New evidence or data might reduce the amount of uncertainty of one type or another, but new evidence generally cannot eliminate altogether any of these types of uncertainty—at least not for any finding of risk that plays an important role in triggering precautions.

⁷ For a more detailed treatment of these five types, see Vern R. Walker, *The Siren Songs of Science: Toward a Taxonomy of Scientific Uncertainty for Decisionmakers*, 23 CONN. L. REV. 567, 574-618 (1991) [hereinafter Walker, *Siren Songs*].

1. Conceptual Uncertainty

Any time scientists select particular events to study and particular concepts to use in describing those events, they place a conceptual structure on the world and frame the way that they gather information about those events. For example, the proposition “inhaling air containing high concentrations of benzene can cause leukemia in people” asserts a causal relationship between certain inhalation events and the development of leukemia. That statement does not refer to other potentially causal factors—for example, genetic, developmental, or environmental factors. Moreover, scientific investigators could describe the same real-world events in an indefinite number of ways, using an indefinite number of variables. An event in a person’s life is not merely a benzene exposure. It can have many exposure-descriptions (for example, “spending time within 100 meters of a high-voltage electricity transmission line”), and it can be part of larger processes (for example, “living in Denver”).

Within the context of science, there are very few constraints on inventing scientific variables and using them as classification categories. Scientists select variables they hope will produce measurements (data) that will prove predictive and explanatory.⁸ Such hypotheses generally extend past work into new areas. But decisions about what to study and how to study it are always made in a pragmatic context. Funding agencies have their own missions, priorities, and preferred methods. Editors of professional journals and their peer reviewers have their own views about what kinds of research are likely to be valuable. Moreover, scientific investigators sometimes design their studies to be instrumental to risk managers, who operate under statutory and political mandates, with settled policies on what counts as an adverse effect or a permissible precautionary measure. Such contexts undoubtedly influence which variables investigators choose to study. However, once investigators select their study variables and gather data using those variables, then any causal conclusions drawn from the study are open to question about whether different or additional variables would have produced different results. The selection of variables can result not only in a lack of knowledge about the variables not studied, but also in inferential error about those variables that are

⁸ For discussions of types of measurement variables, see, e.g., EDWIN E. GHISELLI, ET AL., *MEASUREMENT THEORY FOR THE BEHAVIORAL SCIENCES* (1981); HERMAN J. LOETHER & DONALD G. MCTAVISH, *DESCRIPTIVE AND INFERENTIAL STATISTICS: AN INTRODUCTION* 16–24 (4th ed. 1993).

studied.⁹ The uncertainty created by the selection of variables is what I have called "conceptual uncertainty." Conceptual uncertainty is the potential for error created by using particular variables to describe and study the world. As long as scientific resources and funding are limited, deciding which variables deserve to be studied next is not merely a scientific question, but also reflects decisions about societal needs and values.

2. Measurement Uncertainty

Measurement is the process of classifying individual objects or events into the categories of a variable—that is, the process of generating the data for a scientific study. Measuring individual objects or events incurs the possibility of misclassification. Scientists usually divide measurement uncertainty into two kinds of problems: reliability and validity. A measurement process (as well as the resulting data) is said to be "unreliable" to the extent that repeated measurements of the same object or event by the same measurement process would yield inconsistent results in a random fashion.¹⁰ If a researcher measured or classified the same individual repeatedly, using an unreliable process, the individual would score differently under the variable and in a random pattern. For example, repeat analyses of the same air sample might yield different concentrations of benzene in parts per million, but all those results might fall in a random pattern and largely within one percent of the mean value. For many measurement processes, especially in the physical sciences, reliability studies can determine distributions of error under different sets of circumstances. Reducing the range of random variation in measurements increases the "precision" of the measurements.¹¹ In the behavioral sciences, however, the feasibility of re-testing the same individual is especially difficult, since the measurement process itself might change the behavior of the test subject. Re-taking psychological tests, for example, might produce higher scores simply because taking tests improves the individual's test-taking skills. Despite such methodological problems with certain types of subjects and variables, scientists still employ a fairly clear notion of measurement unreliability, in which

⁹ See discussion of causal uncertainty, *infra* Section II.A.5.

¹⁰ See, e.g., EDWARD G. CARMINES & RICHARD A. ZELLER, RELIABILITY AND VALIDITY ASSESSMENT 11–13 (John L. Sullivan ed., 1979); GHISELLI ET AL., *supra* note 8, at 184, 191.

¹¹ See, e.g., Theodore Peters, Jr. & James O. Westgard, *Evaluation of Methods*, in TEXT-BOOK OF CLINICAL CHEMISTRY 410, 412 (Norbert W. Tietz ed., 1986).

the error is due to random variations in the measurement process itself.

Using a perfectly reliable measurement process that yields exactly the same measurement score for the same individual every time the process is repeated still could leave uncertainty about the "validity" of the measurement. A measurement process is valid to the extent that it measures exactly what it is thought to measure.¹² Measurement processes are invalid when they place an individual object or event in the wrong category of the variable, even if they do so repeatedly and consistently. Validity problems raise external questions about the "accuracy" of the measurement process, as determined by alternative measures of the same variable.¹³ If, for example, a "criterion method" or "reference method" exists that serves as a standard for measuring benzene in air, then a new technology or process for measuring the same variable would be tested for validity against the criterion method.¹⁴ To the extent that the new measurement process produces biased data or systematic error, as compared to the results obtained by the criterion method, then there is a validity problem.

While reliability is a matter of internal consistency (using the same method on the same subject), validity is a matter of external consistency (different methods but the same subjects). Both kinds of measurement uncertainty are about the measurement process, not about true variations in the individual objects or events being measured. Moreover, both kinds of measurement uncertainty require decisions about the level of uncertainty that is acceptable in factfinding. Because every measurement process is to some extent unreliable, the factfinder must decide upon the level of "acceptable imprecision." Because it is usually possible to question the degree of invalidity involved in a measurement process, the factfinder must decide upon the level of "acceptable inaccuracy." The acceptable imprecision and inaccuracy for a space shuttle launch are likely to be different than what is acceptable for flying commercial aircraft or for driving a car. Factfinding about the risk that ought to trigger the taking of precautions likewise varies from case to case. Therefore, finding that the

¹² See, e.g., CARMINES & ZELLER, *supra* note 10, at 12–13; GHISELLI ET AL., *supra* note 8, at 266.

¹³ Peters & Westgard, *supra* note 11, at 412; David H. Kaye & David A. Freedman, *Reference Guide on Statistics*, in FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 83, 103–04 (2d ed. 2000).

¹⁴ See, e.g., Peters & Westgard, *supra* note 11, at 412; Kaye & Freedman, *supra* note 13, at 104.

measurement uncertainty in a risk determination is within acceptable bounds is not a purely scientific decision.

3. Sampling Uncertainty

Scientific data record actual measurements taken on particular objects or events. Scientists often want to generalize beyond the past measurements, however, and they warrant generalizations about objects or events as yet unmeasured. Scientists distinguish between the "sample" (the individuals actually measured, or the data gathered from measuring them) and the "population" (the group that is the subject of the generalization). Making an inference from sample data to a conclusion about the population creates the possibility that error will be introduced because the sample does not adequately represent that population. Whether a generalization is warranted depends in part upon the nature of the relevant variable and the nature of the individual objects or events studied. For example, the results of an occupational health study might warrant generalization to the general population because the biological processes being studied are fairly uniform among people. On the other hand, a worker sample might contain an unrepresentative proportion of healthy subjects, relative to the general population. This could create a "healthy worker" bias in an inference from sample results to the general population.

Scientists prefer to warrant a generalization by the manner in which they draw the sample and analyze the data. If possible, they draw the sample in such a way that they can warrant assigning a probability distribution to all possible samples and statistical results.¹⁵ For example, if the population contains 50% men and 50% women, then there is a certain probability of randomly drawing a sample of 200 people in which there would be 102 men and 98 women. Drawing a scientific or probability sample allows scientists to calculate the mathematical probability of drawing a particular type of sample from a particular type of population. Once a scientific sample has been drawn and those probabilities have been calculated, the sample results can help warrant conclusions about the population itself. For example, suppose a probability sample of 200 people, drawn from a population with an unknown proportion of men and women, contains 150 men and only 50 women. The hypothesis that the population contains 50% men and 50% women might be implausible given

¹⁵ See, e.g., WILLIAM L. HAYS, STATISTICS 224-25 (5th ed. 1994).

this sample. The probability of drawing such a sample from such a population may be extremely low—so low that a scientist would conclude that this hypothesis about the make-up of the population is probably false. The warrant for rejecting this hypothesis as improbable rests on the way the sample was drawn, on the actual sample results, and on probability theory itself.

Statisticians have invented various techniques for warranting inferences about the population on the basis of a probability sample. These techniques include hypothesis testing using P-values, confidence intervals, and statistical power.¹⁶ These techniques characterize the extent of the sampling uncertainty inherent in any inferences from sample to population. They help characterize the potential for error that is created by the fact that the empirical evidence is limited to sample data. Such techniques cannot eliminate the possibility of sampling error, but they can aid in characterizing the degree of sampling uncertainty involved in a generalization. Ultimately, however, every reasonable factfinder must decide what degree of sampling uncertainty is acceptable. For example, one convention among scientists is to consider “statistical significance at the 0.05 level” as acceptable for rejecting an hypothesis about a population.¹⁷ Such a convention reflects decision-making *among scientists*, but is not itself a scientific conclusion. Selecting the level of sampling uncertainty that is acceptable for precautionary purposes is not a purely scientific decision.

4. Modeling Uncertainty

Measurement and sampling uncertainty exist even for data gathered on a single variable. However, generalizations about risk rest upon causal relationships among multiple variables. Scientists use mathematical models to predict values for some variables based on values for other variables. Two important examples of such mathematical models in risk determinations are relative risk and linear regression models.¹⁸ Epidemiologists and public health officials often

¹⁶ See, e.g., Kaye & Freedman, *supra* note 13, at 373–93.

¹⁷ See, e.g., CARL F. CRANOR, *REGULATING TOXIC SUBSTANCES: A PHILOSOPHY OF SCIENCE AND THE LAW* 29–48 (1993); Neil B. Cohen, *Confidence in Probability: Burdens of Persuasion in a World of Imperfect Knowledge*, 60 N.Y.U. L. REV. 385, 409–17 (1985); Michael Cowles & Caroline Davis, *On the Origins of the .05 Level of Statistical Significance*, 37 AM. PSYCHOLOGIST 553 (1982); HAYS, *supra* note 15, at 267–82, 302–03.

¹⁸ Walker, *Siren Songs*, *supra* note 7, at 598–608 (linear regression model); Vern R. Walker, *The Concept of Baseline Risk in Tort Litigation*, 80 KY. L.J. 631, 651–62 (1992) (relative

use relative risk to characterize or predict the excess risk of populations exposed to a hazard.¹⁹ Toxicologists and exposure modelers often use linear regression models to characterize the incremental contributions of multiple hazards (e.g., asbestos exposure and cigarette smoking) to the total risk of an adverse effect (e.g., lung cancer). Scientists use a variety of mathematical models to characterize the quantitative relationships among multiple variables.

Even a very simple mathematical model will illustrate the nature of the uncertainty that use of such models introduces. For example, if a particular clock "runs fast," we can still use it to determine the time of day if we use the right mathematical model. If the watch gains approximately two minutes per hour, then one model might be to multiply two minutes times the number of hours since the clock was last set, and subtract that product from the time shown on the clock. Such a model creates uncertainty in two ways. First, using a different adjustment parameter or "constant" in the formula might yield better results (e.g., 110 seconds instead of two minutes). Second, using a different formula might be more accurate (e.g., multiplying the duration of the previous hourly cycle by 1.02, instead of adding a constant value for each passing hour). Using different parameters or different model formulae will produce different predictions, with different levels of precision and accuracy. Such uncertainties are similar to the issues of reliability and validity discussed above under measurement uncertainty. When the evidence of risk involves mathematical models, the factfinder faces similar decisions about what levels of model reliability and validity are acceptable in particular contexts. These decisions are not purely scientific, but instead require judgments about what potential for error is acceptable given what is at stake.

5. Causal Uncertainty

Even when measurement, sampling, and modeling uncertainties are acceptable, and the evidence warrants that there are probably quantitative associations among events in the real world, an additional potential for error still exists in interpreting the underlying causation that explains such associations. Warranted mathematical models can

risk model); see Michael D. Green et al., *Reference Guide on Epidemiology*, in FEDERAL JUDICIAL CENTER, *supra* note 13, at 348–54 (relative risk and odds ratio in epidemiologic studies); David L. Rubinfeld, *Reference Guide on Multiple Regression*, in FEDERAL JUDICIAL CENTER, *supra* note 13, at 179–227.

¹⁹ See, e.g., Green et al., *supra* note 18, at 348–54.

predict events based on other events, but causal relationships are needed to explain why those events occur. Two types of events can occur or vary together without one causing the other. For example, barometric changes do not cause weather patterns, and clinical symptoms do not usually cause the disease. Even if there exists a real statistical association between events of type A and events of type B, there are numerous causal possibilities: A might cause B; B might cause A; A and B might interact in complicated ways; other types of events might cause both A and B; and so on. One result of conceptual uncertainty (gathering data only on certain variables and not on others) may be that the truly explanatory variables are ignored and the resulting causal theories are inaccurate.²⁰ The causal account of the statistical association may remain largely unknown.

On the other hand, even if there is some evidence that no statistical association exists between two types of events, those events might still be related causally. For example, some unstudied event might counteract or mask the causal action.²¹ The complexities of human metabolism often make it very difficult to determine causal patterns within the human body. The same is true of complex ecosystems. A controlled experiment might manipulate some suspected causal factors in such a way that it unmasks a causal influence, but controlled experiments are not always feasible. For ethical, methodological, and economic reasons, the best available evidence is often epidemiological or the product of field studies; thus, the resulting conclusions about causal action are subject to significant causal uncertainty. Unless the causal system is closed and completely understood, the factfinder must decide, explicitly or implicitly, what level of causal uncertainty is acceptable in the context. Again, such a decision is not purely scientific.

B. Risk Assessment, Scientific Uncertainty, and Science Policies

Since the 1970s, those involved in health, safety, and environmental regulation have developed a set of methodologies called "risk assessment."²² A risk assessment process attempts to determine which

²⁰ See discussion *supra* Section II.A.1.

²¹ See, e.g., JAMES A. DAVIS, *THE LOGIC OF CAUSAL ORDER* 25–27 (1985).

²² See generally NATIONAL RESEARCH COUNCIL, *RISK ASSESSMENT IN THE FEDERAL GOVERNMENT: MANAGING THE PROCESS* (1983) [hereinafter NRC (1983)]; NATIONAL RESEARCH COUNCIL, *SCIENCE AND JUDGMENT IN RISK ASSESSMENT* (1994) [hereinafter NRC (1994)]; 1 *THE PRESIDENTIAL/CONGRESSIONAL COMMISSION ON RISK ASSESSMENT AND*

adverse effects can be caused by exposure to a toxic agent, and the probability that such exposure will lead to such effects.²³ Risk assessment, therefore, divides into two major sub-issues: the toxicity of the causal agent or hazard, and the predicted exposure to that hazard. Toxicity assessment evaluates the qualitative and quantitative aspects of the causal relationship between a dose or level of exposure and a resulting incidence or severity of adverse effect.²⁴ Exposure assessment evaluates the probability, magnitude, duration, and timing of the doses that target organisms might receive through the various pathways of exposure (such as inhalation, ingestion, or dermal absorption).²⁵ Total risk increases as either toxicity or potential exposure increases. Protective strategies might be to decrease toxicity or exposure, or both. Risk assessment combines toxicity and exposure assessments in order to characterize the total risk posed by some condition or event. The purpose of risk assessment is to provide accurate and useful risk characterizations to risk managers, who then can decide what should or will be done to reduce or manage those risks.

A significant amount of scientific uncertainty exists within virtually every risk assessment of any importance. In most cases, there is a well-recognized lack of information both about the toxicity of an agent (i.e., the conditions under which exposure to it can cause adverse effects) and about the likelihood of exposure. Scientists do not have complete toxicity studies for the vast majority of chemical compounds employed today. Even when toxicity studies do exist, significant uncertainties usually remain. For example, even after toxicologists conduct carcinogenicity studies in animals, uncertainties remain about whether the test animals are adequate biological models for humans; whether there is a no-risk threshold somewhere between the high-dose effects observed in the study and the low-dose exposures encountered outside the laboratory; and whether there exists an

RISK MANAGEMENT, FRAMEWORK FOR ENVIRONMENTAL HEALTH RISK MANAGEMENT (1997).

²³ This notion of risk assessment is now incorporated into an international treaty. See Agreement on the Application of Sanitary and Phytosanitary Measures, Apr. 15, 1994, WTO Agreement, Annex 1A, LEGAL INSTRUMENTS--RESULTS OF THE URUGUAY ROUND vol. 1 (1994), available at <http://docsonline.wto.org> [hereinafter SPS Agreement].

²⁴ See, e.g., NRC (1983), *supra* note 22, at 19-27; NRC (1994), *supra* note 22, at 26, 57-66; Commission Directive 93/67, 1993 O.J. (L 227) 9 ("laying down the principles for assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/548/EEC").

²⁵ NRC (1983), *supra* note 22, at 20, 27-28; NRC (1994), *supra* note 22, at 26-27, 43-55; Commission Directive 93/67, *supra* note 24.

adequate margin of safety to protect unusually sensitive people?²⁶ Similarly, even after researchers complete a typical exposure assessment, there may be significant uncertainties about the frequency, magnitude, and duration of individual exposures through diet, occupation, and other scenarios. Risk assessors must make decisions about how to take into account the vast number of remaining uncertainties, even when the available studies are relatively extensive.²⁷ Despite the pervasiveness of scientific uncertainty in risk assessment, the myth of science as a "neutral arbiter" about risk often describes the factfinding process using the neutral-sounding language of "risk assessment." Risk assessments that arrive at findings about risk, however, in the face of scientific uncertainty, cannot be purely scientific.

One important question is *who* will make those decisions that are necessarily inherent but not scientific in nature. One possibility is to let scientists make those decisions, despite the fact that the decisions are not scientific ones. Individual scientists make many decisions about which variables are relevant to a study, how to handle missing data, when to make simplifying assumptions, which mathematical models to employ, and so forth. Groups of scientists sometimes establish conventions for which decisions to make in the face of a commonly encountered problem. The convention about an acceptable level of statistical significance is one such example. Conventions about how to analyze and report experimental results allow subsequent researchers to duplicate previously performed studies, thereby providing a check on the earlier study's results. Modern science became a social endeavor, with powerful epistemic advances, only after methodological norms and conventions created a common enterprise. But decisions by scientists to adopt or adhere to a methodological convention are often merely decisions about how to proceed, not scientifically warranted conclusions about the world. Nevertheless, per-

²⁶ See, e.g., Vern R. Walker, *Keeping the WTO from Becoming the "World Trans-science Organization": Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute*, 31 CORNELL INT'L L.J. 251, 263-66 (1998).

²⁷ One option for risk assessment scientists, of course, is merely to communicate all of the residual uncertainties to others, and to leave the ultimate task of factfinding about risk to those others. I express no opinion here about the advantages and disadvantages of such an approach. The analysis here applies to conclusions or findings about the existence or non-existence of a triggering risk, regardless of whether scientists or non-scientists make those ultimate findings. In addition, the analysis in Section II of this article applies even to findings about causation, which provide the evidentiary grounds for ultimate findings about triggering risks.

haps the needed decision-making could be left to scientists, especially if there were no better alternative.

An alternative has begun to appear, however, within the risk assessment context. When risk assessors encounter identifiable and recurrent instances of scientific uncertainty, governmental decision-makers sometimes establish explicit "science policies" for risk assessors to follow.²⁸ Science policies are decision rules about the way in which risk assessment scientists should proceed when they encounter specified types of uncertainties.²⁹ Science policies direct the choices that risk assessors make from among the scientifically plausible assumptions. What makes uncertainty "scientific" is that scientists might agree on a number of *plausible* accounts, but cannot determine through scientifically accepted methods which of those plausible accounts will ultimately prove to be the correct one. For example, if no data are available for the rate of dermal absorption of a particular chemical, then scientists might agree on a plausible range of dermal absorption values, but might lack the data to narrow that range further. A science policy might then prescribe which of those plausible default values to assume (e.g., a particular default value for the rate of dermal absorption in adults). Another example is that scientists might agree that certain mathematical curves are not plausible candidates for characterizing the dose-response function for a particular chemical, but they might disagree on which of the plausible curves best approximates the dose-response function. A science policy might then direct which curve to employ as a default (e.g., an assumption that there is no "safe" or "threshold" dose).

Such science policies are not themselves scientific conclusions. If scientific evidence were able to resolve the uncertainty, there would be no need for a science policy. Science policies tell the risk assessor how to proceed in a principled way to characterize the total risk, although doing so means reaching conclusions beyond what scientific method would warrant. Science policies establish decision rules to deal with uncertainties for which scientific conventions do not exist. Such policies often pick up where established scientific conventions

²⁸ See, e.g., EPA, Guidelines for Carcinogen Risk Assessment, 51 Fed. Reg. 33,992 (1986); EPA, Proposed Guidelines for Carcinogen Risk Assessment, 61 Fed. Reg. 17,960 (1996); NRC (1983), *supra* note 22; NRC (1994), *supra* note 22; Walker, *supra* note 26, at 258-63.

²⁹ Science policies are also said to be "inference guidelines" or to result in "default assumptions." See NRC (1983), *supra* note 22, at 28-37; NRC (1994), *supra* note 22, at 27, 85-90; EPA, Proposed Guidelines for Carcinogen Risk Assessment, *supra* note 28, at 17,964.

end, and guide risk assessors in situations not covered by established conventions.

But if scientists facing uncertainty make their own assumptions, either on an ad hoc basis or by adopting conventions, why should administrative agencies establish explicit science policies to govern findings about risk, instead of simply "letting scientists do whatever scientists do"? Agencies give a number of rationales for adopting explicit science policies,³⁰ which are mentioned here only briefly. First, explicit science policies help findings about risk to be "transparent." Explicit science policies for commonly encountered situations of uncertainty allow everyone, including scientists, regulators, and the potentially affected public, to distinguish the scientifically warranted inferences from the non-scientific decision-making. As a result, the scientific basis for the risk characterization is distinct from the policy component. Greater transparency creates greater clarity in risk communication, for the acknowledged application of explicit policies by factfinders communicates to risk managers and to the general public the true mix of science and policy within the risk assessment.

Second, explicit science policies enable governmental institutions to distinguish the proper functions of risk assessors and risk managers. For example, a science policy might institute a presumption that if a compound can cause cancer in test animals, then it can also do so in humans. This presumption directs risk assessors to use positive animal data to characterize the extent of risk to humans. Such a policy enables risk assessment scientists to complete risk characterization, without the scientists themselves making the non-scientific decisions underlying those findings. At the same time, the existence of an explicit science policy helps to ensure that risk managers (not scientists) will make the decisions underlying the science policies, and that those decisions will receive appropriate policy justifications. The use of explicit science policies allows all participants and affected parties to distinguish the two activities of scientific inference and non-scientific decision-making, and encourages everyone to evaluate science as science and policy as policy. Scientists and scientific communities seldom attempt to justify their conventions on the basis of policies. Governmental institutions, by contrast, are more likely to recognize default rules as being decisions that require justification—by appeal,

³⁰ See, e.g., EPA, Memorandum: EPA Risk Characterization Program (1995); EPA, Policy for Risk Characterization (1995); EPA, Science Policy Council, Guidance for Risk Characterization (1995); Walker, *supra* note 26, at 260–62.

for example, to the desirability of the consequences, the effectiveness of the rule on the whole, the fairness of the approach, and the equity and efficiency of following the same rule in all similar cases.

Third, making science policies explicit increases the likelihood of uniformity between one finding of risk and another, regardless of differences in the regulatory context. To the extent that all risk assessors make the same default inferences when faced with similar circumstances of uncertainty, potentially affected parties can be more confident that the findings rest on comparable assumptions and decisions. This also allows meaningful comparisons among risk findings and a common basis for setting priorities. Achieving uniformity also serves non-epistemic goals of governance, such as the equitable treatment of potentially affected parties. When decision-makers bridge gaps in knowledge by instituting explicit science policies, they acknowledge that what they are doing is not "pure science" at all, but rather decision-making in the service of regulatory objectives.

III. EXAMPLES OF TRIGGERING FINDINGS ABOUT RISK

This section discusses a variety of findings about risk from the law of the United States, the European Community, and international trade law. These findings are used to trigger or justify the taking of precautions in a variety of areas. These examples are discussed here only briefly, to illustrate the non-scientific decisions involved in making such findings. This discussion does not evaluate the strength of evidence for any particular finding in any particular case.

A. Examples of Food Safety Triggers in the United States

Under the food safety law of the United States, a number of factual situations can trigger precautionary measures.³¹ The Federal Food, Drug, and Cosmetic Act (FFDCA) is the cornerstone statute, and is oriented toward ensuring that food is "safe."³² Any marketed food³³ is subject to enforcement or regulation if it bears or contains a substance that is "poisonous or deleterious" and that substance "may

³¹ For a more complete review of food safety law from the standpoint of taking precautions, see Vern R. Walker, *Some Dangers of Taking Precautions Without Adopting the Precautionary Principle: A Critique of Food Safety Regulation in the United States*, 31 ENVTL. L. REP. 10040 (2001).

³² 21 U.S.C. § 342 (2000).

³³ The word "marketed" here refers to any acts that are prohibited under 21 U.S.C. § 331.

render" the food "injurious to health."³⁴ Moreover, if a substance that is "poisonous or deleterious" is "added to" the food, then the food is subject to enforcement or regulation under certain conditions.³⁵ These triggering predicates require non-scientific decisions about their meaning. First, the statutory words "poisonous or deleterious" and "injurious to health" have narrower meanings than the word "harm" used in Section I above. Some kinds of harm might not serve as triggers under particular sections of the FFDCa, such as economic harm (e.g., food that is too expensive) or aesthetic harm (e.g., food that appears unappetizing).³⁶ Moreover, there is a threshold of *de minimis* "risk" to be determined, at least for natural substances that are not "added to" food. Even if such a substance is "poisonous or deleterious," there must be a finding that its presence may render the food itself "injurious to health." The food cannot be considered adulterated "if the quantity of such substance in such food does not ordinarily render it injurious to health."³⁷ Second, the statutory language displays various modalities, from the subjunctive verb "may render" to the indicative "does not ordinarily render."

The statutory dichotomy between "safe" and "unsafe" takes on more precise meanings in other statutory requirements, administrative rules, and judicial decisions. Congress itself has determined that certain categories of substances in food are "unsafe" unless proven to be safe: pesticide chemical residues,³⁸ food additives,³⁹ color addi-

³⁴ *Id.* § 342(a)(1).

³⁵ *Id.* §§ 342(a)(2)(A), 346.

³⁶ See, e.g., *United States v. 449 Cases Containing "Tomato Paste,"* 212 F.2d 567, 569-73 (2d Cir. 1954) (surveying various factual predicates that serve as regulatory triggers under the FFDCa); *United States v. 24 Cases, More or Less,* 87 F. Supp. 826, 828 (D. Me. 1949) (holding, in a libel in rem proceeding relating to cases of canned herring roe, that "in order for a product to be subject to condemnation as unfit for food, on account of it's [sic] tough and rubbery consistency, the product must be proved to be so tough and rubbery that the average, normal person, under ordinary conditions, would not chew and swallow it").

³⁷ 21 U.S.C. § 342(a)(1); 21 C.F.R. §§ 109.3(c), (d) (2000). Cf. PETER BARTON HUTT & RICHARD A. MERRILL, *FOOD AND DRUG LAW, CASES AND MATERIALS* 948-50 (2d ed. 1991) (carcinogenic natural foods and ingredients).

³⁸ 21 U.S.C. §§ 321(q), 342(a)(2)(B), 346a(a)(1). The statute declares any pesticide chemical residue in or on a food to be "unsafe" unless its use conforms to a regulatory tolerance or an exemption. *Id.* §§ 342(a)(2)(B), 346a(a).

³⁹ 21 U.S.C. §§ 321(s), 342(a)(2)(C)(i), 348(a). The statute declares any food additive (other than a food contact substance) to be "unsafe" unless its use conforms to an exemption or a use regulation. *Id.* §§ 342(a)(2)(C)(i), 348(a).

A precondition for being a "food additive" is that the substance not be "generally recognized, among experts qualified by scientific training and experience to evaluate its safety, as having been adequately shown . . . to be safe under the conditions of its intended

tives,⁴⁰ and new animal drugs or conversion products of new animal drugs.⁴¹ Congress has determined that there is an unacceptable risk in putting such substances into the food chain without pre-marketing approval.⁴² The pre-marketing requirements generally consist of producing required types of evidence of safety and persuading the relevant agency that any existing risk of harm is acceptable. For example, in order for the Food and Drug Administration (FDA) to approve color additives as "safe," the agency requires "convincing evidence that establishes with reasonable certainty that no harm will result from the intended use."⁴³ The administrative definition of "safe" for food additives is similar: "safe . . . means that there is a reasonable certainty in the minds of competent scientists that the substance is not harmful under the intended conditions of use."⁴⁴ The wording of these findings displays the various elements that require non-scientific decisions: the meaning of the categorical predicate "harm," and the risk of such harm posed by "intended use"; the seemingly conservative modifiers "no harm" and "not harmful"; and the probabilistic modality indicated by "reasonable certainty." Moreover, the agency clearly

use"—that is, that the substance is not "generally recognized as safe" (GRAS). *Id.* § 321(s). Being GRAS is therefore based on "scientific procedures"—unless the substance was used in food prior to 1958, in which case an alternative basis is "experience based on common use in food." *Id.* If a substance is GRAS, then it is not a food additive and does not trigger the pre-marketing approval required for food additives. 21 C.F.R. § 170.30(g) (2000).

⁴⁰ 21 U.S.C. §§ 321(t), 342(c), 379e(a). The statute declares any color additive to be "unsafe" unless its use conforms to a use regulation or an exemption. *Id.* §§ 342(c), 379e(a).

⁴¹ *Id.* §§ 321(v), 342(a)(2)(C)(ii), 360b(a). The statute declares any new animal drug to be "unsafe" unless its use conforms to an approved application. *Id.* §§ 342(a)(2)(C)(ii), 360b(a)(1). Excluded from this category are drugs "generally recognized . . . as safe and effective" by qualified experts and other "grandfathered" drugs. *See id.* § 321(v).

⁴² Other pre-marketing triggers also exist, such as the inspection-and-condemnation provisions for meat and meat food products, poultry products, and egg products. *Id.* §§ 601 ff, 610(c)(2), 451 ff, 458(a)(2)(B), 1031 ff, 1037(b)(2). For such products, inspection measures are triggered categorically for the relevant animals or products, or for establishments that process them. *See, e.g., id.* §§ 455, 603, 604, 608, 1034. If the required inspection results in an administrative determination that a particular product is "adulterated," then that product can be condemned and destroyed. *See, e.g., id.* §§ 604, 672, 673, 455(c), 467a, 467b, 1034(c), 1048, 1049. The relevant agency also has the authority to prosecute judicial proceedings, and the court may condemn and seize any covered product that is found to be adulterated and capable of being used as human food. *See, e.g., id.* §§ 610(c)(1), 673(a)(1)(B), 458(a)(2)(A), 467b(a)(1)(B). "Adulteration" is generally defined as bearing or containing any poisonous or deleterious substance that may render the product either injurious to health or unfit for human food. *See, e.g., id.* §§ 601(m), 453(g), 1033(a).

⁴³ 21 C.F.R. § 70.3(i) (2000).

⁴⁴ *Id.* § 170.3(i).

opens the way for placing requirements on supporting evidence through the phrases "convincing evidence" and "reasonable certainty in the minds of competent scientists." The language suggests that such findings are purely scientific in nature, when demonstrably they are not.

The non-scientific nature of such findings becomes more clear when factfinders decide that low levels of risk are so acceptable that they determine the food to be "safe." Under the traditional approach for adverse health effects other than cancer, when scientific evidence from animal studies indicates that there exists a threshold dose for adverse effects, risk managers have adopted policies for acceptable margins of safety.⁴⁵ Generally, they have decided that a factor of 100 provides a margin of safety that makes any remaining risk acceptable. That is, a substance that is harmful to animals at doses only above a "no-adverse-effect level" has been considered "safe" for humans (in the statutory meaning of "safe") at doses less than 1/100th of that animal no-adverse-effect level.⁴⁶ When the Environmental Protection Agency (EPA) assesses the risk of pesticide chemical residues, however, Congress requires the EPA to adopt "an additional tenfold margin of safety" specifically for infants and children.⁴⁷ As a result, the margins of safety normally applied to non-carcinogenic effects studied with animal data may be different for pesticide residues (1/1000) than they are for food or color additives (1/100), at least when infants and children might be exposed.⁴⁸ Using a specific safety factor for a particular regulatory category is a risk management decision and certainly not a conclusion of pure science.

⁴⁵ See Final OECD-U.S. Summary Precaution Annex, *Precaution in U.S. Food Safety Decisionmaking: Annex II to the U.S. National Food Safety System Paper* (Mar. 2000), §§ B.1, B.3 & ¶ 82 [hereinafter Precaution Annex].

⁴⁶ The normally applied safety factor for a color additive is "1/100th of the maximum no-effect level for the most susceptible experimental animals tested." 21 C.F.R. § 70.40 (2000). The tolerance for a food additive usually will not exceed "1/100th of the maximum amount demonstrated to be without harm to experimental animals." *Id.* § 170.22. The requirement to consider "safety factors" for food and color additives is statutory. 21 U.S.C. §§ 348(c)(5)(C), 379e(b)(5)(A)(iii). For similar use of a 100-fold safety factor by EPA in setting tolerance levels for pesticide residues, see H.R. REP. NO. 104-669, pt. 2, at 41 (1996).

⁴⁷ 21 U.S.C. § 346a(b)(2)(C). EPA is authorized to "use a different margin of safety for the pesticide chemical residue only if, on the basis of reliable data, such margin will be safe for infants and children." *Id.*

⁴⁸ Compare note 46 *supra* with 21 U.S.C. § 346a(b)(2)(C); see H.R. REP. NO. 104-669, pt. 2, at 41, 43 (1996); Frank B. Cross, *The Consequences of Consensus: Dangerous Compromises of the Food Quality Protection Act*, 75 WASH. U.L.Q. 1155, 1166-67 (1997).

The regulation of carcinogens in U.S. food law provides further examples of non-scientific decision-making at both the legislative and administrative levels. Congress has enacted three "Delaney Clauses" that create a "*per se* risk management law"⁴⁹ for carcinogenic substances in three categories: food additives,⁵⁰ color additives,⁵¹ and new animal drugs.⁵² If a food additive, for example, is found to induce cancer when ingested by animals, then the FDA has no discretion to approve that additive as safe.⁵³ During the 1980s, the FDA and the EPA tried to adopt more relaxed interpretations of the Delaney Clauses, but with limited success.⁵⁴ The FDA implemented a policy that a food or color additive containing a carcinogenic chemical "constituent" does not necessarily trigger the Delaney Clause, and the safety of the additive as a whole is determined under the general safety clause.⁵⁵ However, courts generally considered the language and intent of Congress in the Delaney Clauses to be clear and refused to allow relaxed administrative interpretations.⁵⁶ In 1996, partly as a result of a judicial decision involving pesticide residues,⁵⁷ Congress changed the level of protection for carcinogenic pesticide chemical

⁴⁹ Precaution Annex, *supra* note 45, at ¶¶ 87-92.

⁵⁰ 21 U.S.C. § 348(c)(3)(A).

⁵¹ *Id.* § 379e(b)(5)(B).

⁵² *Id.* § 360b(d)(1)(I).

⁵³ See Precaution Annex, *supra* note 45, ¶¶ 87, 89. Because a food ingredient that is GRAS does not fall within the definition of "food additive," it is not technically subject to the Delaney Clause. HUTT & MERRILL, *supra* note 37, at 338. Moreover, the Delaney Clauses are not triggered unless the agency makes several determinations that normally involve science policies. See, e.g., 21 C.F.R. § 70.50 (2000) (color additives).

⁵⁴ Congress can, of course, exempt food products from the Delaney Clause prohibitions, as it did in the case of saccharin. See HUTT & MERRILL, *supra* note 37, at 922-28.

⁵⁵ See FDA, Policy for Regulating Carcinogenic Chemicals in Food and Color Additives, 47 Fed. Reg. 14,464 (Food & Drug Admin., Advance Notice of Proposed Rulemaking, Apr. 2, 1982); Scott v. FDA, 728 F.2d 322, 325 (6th Cir. 1984) (upholding FDA's determination that the color additive as a whole did not cause cancer in test animals, although the p-toluidine present in minute quantities is carcinogenic when tested separately). Between the Scott decision and 1990, FDA approved more than thirty color additives and indirect food additives containing a variety of trace carcinogenic constituents. HUTT & MERRILL, *supra* note 37, at 922.

⁵⁶ Pub. Citizen v. Young, 831 F.2d 1108, 1113 (D.C. Cir. 1987), *cert. denied sub nom.* Cosmetic, Toiletry and Fragrance Ass'n v. Pub. Citizen, 485 U.S. 1006 (1988) (rejecting FDA's *de minimis* re-interpretation of the color additives Delaney Clause); Les v. Reilly, 968 F.2d 985, 990 (9th Cir. 1992), *cert. denied sub nom.* Nat'l Agric. Chems. Ass'n v. Les, 507 U.S. 950 (1993) (rejecting EPA's *de minimis* re-interpretation of the food additives Delaney Clause, in a case involving carcinogenic pesticides that concentrated during processing and therefore became food additives in processed food under the FFDCAs as it was then written).

⁵⁷ Les v. Reilly, *supra* note 56.

residues from the zero tolerance of the Delaney Clauses to "a reasonable certainty that no harm will result."⁵⁸ This statutory standard appears to track the FDA's administrative definitions of "safe" for color additives and food additives, and the logical analysis above of the administrative definition applies to the statutory wording as well.⁵⁹ In the case of carcinogens, Congress expected this level of protection to be a lifetime risk no greater than one in one million, calculated using conservative assumptions.⁶⁰ When qualitative evidence shows that a substance can cause cancer in animals at high doses, the decision that a specific low dose is "safe" for humans (using the statutory meaning of "safe") is a decision of risk management, not a conclusion of pure science.

B. Examples of Risk Triggers in the European Community

Two recent decisions by the European Court of First Instance provide examples of risk triggers within European Community law. The cases involved judicial review of withdrawals of Community authorization for two antibiotics as additives in animal feed: virginiamycin, at issue in *Pfizer Animal Health SA v. Council*,⁶¹ and bacitracin zinc, at issue in *Alpharma Inc. v. Council*.⁶² These decisions are part of the European Community's evolving law on its precautionary princi-

⁵⁸ See Food Quality Protection Act of 1996, Pub. L. no. 104-170, 110 Stat. 1489, 1514-35, § 405 [hereinafter FQPA]; 21 U.S.C. § 346a(b)(2)(A)(ii); H.R. REP. NO. 104-669, pt. 1 at 36, pt. 2 at 40 (1996); Cross, *supra* note 48, at 1160-62. Under the prior law, pesticide residues that concentrated in food processing triggered the regulatory provisions for food additives, including the food additive Delaney Clause; see *Les v. Reilly*, *supra* note 56.

⁵⁹ See *supra* text accompanying notes 43-44; Cross, *supra* note 48, at 1162-63.

⁶⁰ See H.R. REP. NO. 104-669, pt. 2, at 41 (1996), which states:

In the case of a nonthreshold effect which can be assessed through quantitative risk assessment, such as a cancer effect, the Committee expects, based on its understanding of current EPA practice, that a tolerance will be considered to provide a "reasonable certainty of no harm" if any increase in lifetime risk, based on quantitative risk assessment using conservative assumptions, will be no greater than "negligible." It is the Committee's understanding that, under current EPA practice, . . . EPA interprets a negligible risk to be a one-in-a-million lifetime risk. The Committee expects the Administrator to continue to follow this interpretation.

Id. See Cross, *supra* note 48, at 1164-65. This quantitative level of protection of one in one million is precisely the risk that FDA had announced to be negligible or *de minimis*. See *c.g.*, HUFF & MERRILL, *supra* note 37, at 899-904; Pub. Citizen v. Bowen, 833 F.2d 364, 365 (D.C. Cir. 1987).

⁶¹ Case T-13/99, *Pfizer Animal Health SA/NV v. Council*, 2002 E.C.R. II __ (2002).

⁶² Case T-70/99, *Alpharma, Inc. v. Council*, 2002 E.C.R. II __ (2002).

ple and contain one approach to findings that trigger application of this principle.⁶³ Therefore, while the two cases deal with withdrawals of authorization for additives in animal feed, the interpretive approach probably applies to a much wider array of legal contexts.

Both the *Pfizer* and *Alpharma* cases reviewed a regulation adopted by the Council of the European Union⁶⁴ that withdrew authorization pursuant to Council Directive 70/524/EEC (as amended)⁶⁵ which is founded in turn on Community authority over a common agricultural policy.⁶⁶ Article 3 of that Directive requires that “no additive may be put into circulation unless a Community authorization has been granted” by the European Commission.⁶⁷ The Council found additives in animal feed to pose sufficient generic risk, such that marketing additives (“putting them into circulation”) is unlawful without pre-marketing approval. Article 3a states five necessary conditions that must be met before Community authorization will be given. The following two necessary conditions are of interest here:

(b) taking account of the conditions of use, it [the additive] does not adversely affect human or animal health or the environment, nor harm the consumer by impairing the characteristics of animal products; . . .

(e) for serious reasons concerning human or animal health its use must not be restricted to medical or veterinary purposes.

Once authorization for an additive has been given, various procedures govern withdrawing that authorization—provided either by Ar-

⁶³ For discussions of that evolving law, see, e.g., Theofanis Christoforou, *The Origins, Content and Role of the Precautionary Principle in European Community Law*, in *LE PRINCIPE DE PRÉCAUTION: ASPECTS DE DROIT INTERNATIONAL ET COMMUNAUTAIRE* 205–30 (Charles Leben & Joe Verhoeven eds., 2002); Theofanis Christoforou, *The Precautionary Principle, Risk Assessment, and the Comparative Role of Science in the European Community and the United States Legal Systems*, in *GREEN GIANTS? ENVIRONMENTAL POLICY OF THE UNITED STATES AND THE EUROPEAN UNION* (N. Vig and M. Faure eds.) (forthcoming 2003).

⁶⁴ Council Regulation 2821/98, 1998 O.J. (L 351).

⁶⁵ *Alpharma*, 2002 E.C.R. II __, ¶¶ 3–6, 17–20, 83, 93–95; *Pfizer*, 2002 E.C.R. II __, ¶¶ 3–6, 17–19, 91, 101–03.

⁶⁶ TREATY ESTABLISHING THE EUROPEAN COMMUNITY, art. 37, Nov 10, 1997 O.J. (C 340) 3 (1997) [hereinafter EC TREATY]; *Alpharma*, 2002 E.C.R. II __, ¶ 132; *Pfizer*, 2002 E.C.R. II __, ¶ 111.

⁶⁷ Council Directive 70/524/EEC, art. 3, 1970 O.J. (L 270).

articles 11 and 24 (applicable in the *Pfizer* case)⁶⁸ or by Article 23 (applicable in the *Alpharma* case).⁶⁹

The two cases involved similar findings for triggering a ban (i.e., for withdrawal of authorization) and a similar evidentiary warrant for those findings. At issue was whether use of these two antibiotics in animal feed increases the risk of microbial resistance to antibiotics used in humans. In the *Pfizer* case, for example, Pfizer did not dispute that, in principle, precautionary measures could be triggered by a warranted finding that the use of virginiamycin as a growth promoter in animal feed “involves a risk of a transfer of antimicrobial resistance from animals to humans”⁷⁰ Pfizer argued that the scientific evidence available to the Community institutions did not warrant such a finding.⁷¹ The European Commission had found, however, that withdrawal of authorization was necessary to ensure the protection of human health.⁷² The Council determined that there was a “risk that the effectiveness of certain human medicinal products might be reduced or even eliminated as a result of the use of virginiamycin.”⁷³ After reviewing the record, the Court held that there was adequate scientific evidence available at the time the regulation was adopted to support the Council’s finding “that the use of virginiamycin as an additive in feedingstuffs entailed a risk to human health.”⁷⁴

These cases illustrate the role of non-scientific decision-making in finding a risk that triggers taking precautions. Directive 70/524/EEC permits withdrawal of authorization for an additive if its use “constitutes a danger to . . . human health” and withdrawal is “necessary . . . to ensure the protection of human . . . health.”⁷⁵ The Court interpreted this as meaning a “risk” to human health and defined “risk” as “the possibility that the use of [the additive] will give rise to adverse effects on human health”⁷⁶ The relevant “harm” is therefore broadly defined as being any adverse effect on health. Moreover, a “risk” is a “possibility” that use can cause an adverse effect. The Court elsewhere called this a “possible link” between use and adverse ef-

⁶⁸ Pfizer, 2002 E.C.R. II __, ¶¶ 91, 101–03, 112.

⁶⁹ Alpharma, 2002 E.C.R. II __, ¶¶ 83, 93–95.

⁷⁰ Pfizer, 2002 E.C.R. II __, ¶ 126.

⁷¹ *Id.* ¶ 127.

⁷² *Id.* ¶ 112; Council Directive 70/524/EEC, *supra* note 67, art. 11(3).

⁷³ Pfizer, 2002 E.C.R. II __, ¶ 112.

⁷⁴ *Id.* ¶ 401; *see id.* ¶¶ 315–22, 359–61, 393.

⁷⁵ *See* Council Directive 70/524/EEC, *supra* note 67, art. 11(1), (3).

⁷⁶ Pfizer, 2002 E.C.R. II __, ¶ 138.

fect,⁷⁷ or a capability of causing an adverse effect.⁷⁸ As for the modifier on the required risk, any finding of "a risk to human health" will do.⁷⁹ As for the modality of the finding, the risk cannot be "purely hypothetical" or "mere conjecture,"⁸⁰ but it does not have to be "fully apparent"⁸¹ or "fully demonstrated."⁸² Finally, the Court gave considerable deference to the Council and Commission on when the available evidence is "adequate."⁸³ The scientific evidence does not need to be of the highest methodological quality,⁸⁴ specific to actual antibiotic use as an additive,⁸⁵ complete,⁸⁶ or conclusive.⁸⁷ With such low thresholds on virtually all of the logical aspects of a triggering finding of risk, it is no surprise that the Court upheld taking the very conservative precaution of banning the previously authorized antibiotics.

This approach to a finding of risk is consistent with the approach taken by the European Commission in its "Communication on the Precautionary Principle."⁸⁸ In that Communication, the Commission gave the following informal statement of a typical triggering situation:

[w]hen there are reasonable grounds for concern that potential hazards may affect the environment or human, animal or plant health, and when at the same time the available data preclude a detailed risk evaluation, the precautionary

⁷⁷ *Id.* ¶¶ 312–24, 342–401.

⁷⁸ *Id.* ¶¶ 147, 165.

⁷⁹ *Id.* ¶ 401.

⁸⁰ *Id.* ¶ 143.

⁸¹ Pfizer, 2002 E.C.R. II ___, ¶ 140.

⁸² *Id.* ¶¶ 139, 141, 146, 360, 379, 384–87, 388.

⁸³ *Id.* ¶¶ 144, 161, 168, 323, 389, 393. The following finding by the Court suggests a large number of focal points for non-scientific decisions:

[T]he Court finds that the Community institutions did not exceed the bounds of the discretion conferred on them by the Treaty when they took the view that the various experiments and observations referred to in . . . the contested regulation were not mere conjecture but amounted to sufficiently reliable and cogent scientific evidence for them to conclude that there was a proper scientific basis for a possible link between the use of virginiamycin as an additive in feedingstuffs and the development of streptogramin resistance in humans.

Id. ¶ 389.

⁸⁴ *Id.* ¶¶ 371–76, 379–80, 382–83, 388.

⁸⁵ Pfizer, 2002 E.C.R. II ___, ¶¶ 354, 370.

⁸⁶ *Id.* ¶¶ 160–62, 382–83.

⁸⁷ *Id.* ¶¶ 382–83.

⁸⁸ Communication from the Commission on the Precautionary Principle, COM(2000) 1, 2–4 [hereinafter Communication].

principle has been politically accepted as a risk management strategy in several fields.⁸⁹

This statement displays the various non-scientific aspects discussed in Sections I and II of this article. It allows for a variety of meanings for “harm” (using the phrase “may affect the *environment* or human, animal or plant *health*”) and a minimal meaning for “risk” (“*concern* that *potential* hazards may affect”), as well as a very weak modality (“*may* affect”). It requires “reasonable grounds” in the available scientific evidence, but recognizes that the evidence may not be the best possible evidence (“the available data preclude a detailed risk evaluation”). It leaves open questions about how specific, adequate, or methodologically sound that evidence must be. The *Pfizer* and *Alpharma* cases show that finding a triggering risk involves non-scientific decisions about acceptable levels of uncertainty about risk.⁹⁰ Although the Commission’s Communication admits that the potential consequences of inaction in the face of risk are factors in triggering recourse to the precautionary principle,⁹¹ it stops short of acknowledging the considerable role of non-scientific decision-making in determining risk.⁹² Nevertheless, from invocations of the precautionary principle in practice, it is clear that non-scientific decision-making plays an essential role in finding a triggering risk.

C. Examples of Risk Triggers in World Trade Organization Disputes

Decisions by international tribunals adjudicating trade disputes further illustrate triggering findings of risk. The Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement),⁹³ administered by the WTO, requires WTO members to “ensure that any sanitary or phytosanitary measure is applied only to the extent necessary to protect human, animal or plant life or health, is

⁸⁹ *Id.* at 9. Another statement of trigger in the Communication is: “The precautionary principle is relevant only in the event of a potential risk, even if this risk cannot be fully demonstrated or quantified or its effects determined because of the insufficiency or inconclusive nature of the scientific data.” *Id.* at 13.

⁹⁰ In addition to the discussion above, see *Pfizer*, 2002 E.C.R. II __, ¶ 153.

⁹¹ There was the possibility of irreversible harm at stake in *Pfizer* and *Alpharma*. See *Pfizer*, 2002 E.C.R. II __, ¶ 334.

⁹² The Commission acknowledges that in the face of scientific uncertainty, risk assessment itself can take a “prudential approach” and incorporate prudential aspects, but seems ambivalent about calling this risk management. See Communication, *supra* note 88, at 12. The better approach is to state clearly that science policies for risk assessors are risk management decisions and that findings of risk are not purely scientific.

⁹³ SPS Agreement, *supra* note 23.

based on scientific principles and is not maintained without sufficient scientific evidence.”⁹⁴ In addition, “[m]embers shall ensure that their sanitary or phytosanitary measures are based on an assessment, as appropriate to the circumstances, of the risks to human, animal or plant life or health, taking into account risk assessment techniques developed by the relevant international organizations.”⁹⁵ The phrase “sanitary or phytosanitary measure” refers to any governmental action that is designed to protect against any “risks” that are covered by the SPS Agreement.⁹⁶ Taken together, these provisions require, among other things, that a warranted finding of a relevant risk is required in order to justify the application of any precautionary measure covered by the Agreement.

The SPS Agreement clearly distinguishes between assessing a risk and deciding what to do about a risk once it is found, although the Agreement itself refers only to “risk assessment” and not to “risk management.”⁹⁷ Under the Agreement, a WTO member is entitled to select any level of protection that it considers “appropriate” for its territory and can establish protective measures to achieve that level of protection.⁹⁸ This is another way of saying that each member can decide what level of risk is acceptable.⁹⁹ Such decisions are sovereign acts of government, at least as long as they are internally consistent and employ only justifiable distinctions.¹⁰⁰ The focus of this article, however, is not on the range of management options permitted under the SPS Agreement, but on the non-scientific aspects of risk findings that can trigger legitimate management precautions.

The Meat Hormones case involved a dispute under the SPS Agreement over a European ban on imports of meat and meat products derived from cattle to which certain hormones had been admin-

⁹⁴ *Id.* art. 2.2.

⁹⁵ *Id.* art. 5.1.

⁹⁶ *Id.* annex A.1. For example, a “sanitary or phytosanitary measure” might be a governmental measure applied “to protect human or animal life or health within the territory of the Member from risks arising from additives, contaminants, toxins or disease-causing organisms in foods, beverages or feedstuffs.” *Id.* annex A.1(b). Thus, the SPS Agreement applies to health risks from food and animal feed.

⁹⁷ WTO, Appellate Body, Report on EC Measures Concerning Meat and Meat Products (Hormones), WT/DS26/AB/R & WT/DS48/AB/R, ¶ 181 (Jan. 16, 1998) [hereinafter *Hormones Report*].

⁹⁸ SPS Agreement, *supra* note 23, pmble., arts. 2.1, 3.3, 5, annex A.5.

⁹⁹ *See id.* annex A.5; Communication, *supra* note 88, § 5.2.1.

¹⁰⁰ *See, e.g.*, SPS Agreement, *supra* note 23, art. 5.5, annex A.5; Walker, *supra* note 26, at 268–71.

istered for growth promotion purposes.¹⁰¹ The legitimacy of such a ban under the Agreement depends upon a finding that such products pose a relevant risk of harm. Under the SPS Agreement, the meaning of harm is broad, covering “human or animal life or health.”¹⁰² The covered “risks” in the case were those “arising from additives, contaminants, toxins or disease-causing organisms in foods, beverages or feedstuffs.”¹⁰³ The Appellate Body held that the risk assessment required under the Agreement needs to address the “potential for adverse effects,” and that “potential” here means “possibility,” not “probability.”¹⁰⁴ There is no minimum quantitative threshold of risk required, and finding a risk to exist is consistent with deciding to adopt a level of protection of “zero risk.”¹⁰⁵ Moreover, the modality of the finding is that the risk must be “ascertainable”: not so uncertain as to be “theoretical,” but not so certain as to provide “absolute certainty.”¹⁰⁶ As for evidentiary warrant, the Appellate Body held that the precautionary measure must be “sufficiently supported or reasonably warranted” by the risk assessment.¹⁰⁷ In addition, the supporting scientific evidence must be “sufficiently specific” to the risk posed by “the residues of those hormones found in meat derived from cattle to which the hormones had been administered for growth promotion purposes,”¹⁰⁸ although it need not be quantitative¹⁰⁹ or represent a consensus view among scientists.¹¹⁰

A second case under the SPS Agreement, the Australian Salmon case, involved a dispute over an Australian prohibition on the importation of fresh, chilled, or frozen salmon.¹¹¹ The relevant “harm” at issue was the spread of pests or disease within Australia’s territory, and the biological and economic consequences.¹¹² As decided in the Meat Hormones case above, any risk found must be “ascertainable,” not

¹⁰¹ Hormones Report, *supra* note 97, ¶ 2.

¹⁰² SPS Agreement, *supra* note 23, annex A.1(b).

¹⁰³ *Id.*

¹⁰⁴ Hormones Report, *supra* note 97, ¶¶ 182–84; SPS Agreement, *supra* note 23, annex A.4.

¹⁰⁵ Hormones Report, *supra* note 97, ¶ 186; WTO, Appellate Body, Report on Australia—Measures Affecting Importation of Salmon, WT/DS18/AB/R, ¶ 125 (Oct. 20, 1998) [hereinafter Salmon Report].

¹⁰⁶ Hormones Report, *supra* note 97, ¶ 186.

¹⁰⁷ *Id.* ¶¶ 186, 193.

¹⁰⁸ *Id.* ¶¶ 187, 198–201.

¹⁰⁹ *Id.* ¶ 187.

¹¹⁰ *See id.* ¶ 194.

¹¹¹ Salmon Report, *supra* note 105, ¶ 1.

¹¹² *Id.* ¶¶ 112, 121; SPS Agreement, *supra* note 23, annex A.1(a).

merely "theoretical," and the measure must be reasonably based on a risk assessment.¹¹³ Under the circumstances of the Australian Salmon case, however, an adequate risk assessment must evaluate the "likelihood" or "probability" of the entry, establishment, or spread of disease, not merely the "possibility" of entry, establishment, or spread.¹¹⁴ Such evidence can be either quantitative or qualitative,¹¹⁵ and need not be complete.¹¹⁶

As this brief discussion demonstrates, the interpretation of the SPS Agreement by the Appellate Body illustrates all of the points of non-scientific decision-making discussed in this article. Non-scientific decisions are inherent in the findings about risk needed to justify precautionary measures under the SPS Agreement.

CONCLUSION

There is an unfortunate myth that science can serve as a "neutral arbiter" for triggering precautionary measures. Numerous non-scientific decisions are necessarily involved in both making and warranting findings that a triggering risk exists. Making a finding of risk involves decisions about the meaning of "risk of harm," about the meaning of any qualitative or quantitative modifiers, and about the truth modality of (or degree of confidence in) the finding as a whole. Moreover, every determination that the available scientific evidence warrants a finding of risk involves decisions about the acceptable degree of various types of uncertainty: conceptual uncertainty, measurement uncertainty, sampling uncertainty, modeling uncertainty, and causal uncertainty. All of these decisions may be made case-by-case, or by applying decision rules (science policies) to similar types of cases. This article has provided a variety of examples from the food safety law of the United States, from recent animal feed cases in the European Community, and from Appellate Body decisions in WTO trade disputes. If the myth of science as a "neutral arbiter" for triggering precautions were harmless, indulging in it might not be dangerous. But the myth is not harmless if it leads to the wrong people making non-scientific decisions through the wrong processes and on the wrong bases. Finding a risk that triggers precautions cannot be a purely scientific act, notwithstanding the myth that a "value-neutral" science can do so.

¹¹³ Salmon Report, *supra* note 105, ¶ 125.

¹¹⁴ *Id.* ¶ 123.

¹¹⁵ *Id.* ¶ 125.

¹¹⁶ *Id.* ¶ 130.