

1998

Keeping the WTO from Becoming the "World Trans-science Organization": Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute

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, *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 (1998)

Available at: https://scholarlycommons.law.hofstra.edu/faculty_scholarship/548

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.

Keeping the WTO from Becoming the “World Trans-science Organization”: Scientific Uncertainty, Science Policy, and Factfinding in the Growth Hormones Dispute

Vern R. Walker*

Introduction	252
I. Obligations of Members Under the SPS Agreement.....	255
A. Scientific Uncertainty and Science Policy Within Risk Assessment	256
1. <i>The Nature of Risk Assessment</i>	256
2. <i>The Elements of Risk Assessment</i>	263
a. Hazard Identification	263
b. Dose-Response Assessment	264
c. Exposure Assessment.....	266
d. Risk Characterization.....	266
B. Risk Management	267
1. <i>Selecting a Level of Protection</i>	268
2. <i>Establishing Measures to Achieve the Selected Level of Protection</i>	271
3. <i>Accepting Alternative Measures as Equivalent</i>	272
C. Harmonizing Sanitary Measures Using International Standards.....	272
II. WTO Review of Sanitary Measures Established by Members	277
A. The Proper Substantive Issues Before a Panel	278
1. <i>Reviewing Risk Assessment Determinations</i>	279
2. <i>Reviewing Risk Management Decisions</i>	285
B. The Evidence Available to a Panel	287
C. The Standard and Burden of Proof	290
III. Factfinding in the Growth Hormones Proceeding	296
A. The Proper Substantive Issues Before the Hormones Panel	296

* Professor of Law, Hofstra University. Ph.D., 1975, University of Notre Dame; J.D., 1980, Yale University. The author served as a consultant to the European Commission in the appeal of the panel reports in the growth hormones dispute discussed in this Article, but the views expressed here are not necessarily those of the European Communities.

1. Content of Findings on Risk Assessment	
Determinations	296
2. Content of Findings on Risk Management	
Determinations	305
B. The Evidence Properly Available to the Hormones Panel	
.....	310
C. The Standard and Burden of Proof Employed by the	
Panel	311
Conclusion	319

Introduction

In the time since the scientist Alvin Weinberg first introduced the term "trans-science,"¹ we have come to view the interface between science and regulation as part fact, part policy, and part decision-making. The factual aspect derives from the scientific evidence, which makes some policies more reasonable than others and some decisions more effective or more efficient than others. The policy aspect is reflected in the determination to base decisions upon the best available scientific evidence, but to bridge any remaining gaps in our scientific knowledge with default inference rules based on non-scientific considerations. The decisional aspect stems from the responsibility of regulatory government to act expeditiously, often despite the fact that our scientific knowledge of consequences is incomplete and our policies are vague and conflicting. In the last half of this century, American administrative law has adjusted to the notion that health, safety and environmental regulation is entrusted primarily to agencies that operate at this fact-policy-decisional interface between science and regulation.

With the advent of the Uruguay Round of Multilateral Trade Agreements,² and in particular the Agreement on Sanitary and Phytosanitary

1. Alvin M. Weinberg, *Science and Trans-Science*, 10 MINERVA 209 (1972) (defining "trans-science" questions as "questions which can be asked of science and yet which cannot be answered by science," although "they are, epistemologically speaking, questions of fact . . ."). See Thomas O. McGarity, *Substantive and Procedural Discretion in Administrative Resolution of Science Policy Questions: Regulating Carcinogens in EPA and OSHA*, 67 GEO. L.J. 729, 732-47 (1979) (citing as a "perfect example of a trans-scientific issue" the extrapolation from high-dose carcinogenic effects to low doses, and arguing that science policy questions "lie on a spectrum that ranges from issues of pure scientific fact to issues of pure policy."). In this Article, I do not rely on any particular set of subcategories of trans-science or science policy issues. It is sufficient for my purposes to contrast science and science policy, see *infra* Part I, and to consider trans-science issues as particularly compelling examples of determinations requiring the use of science policies.

For general discussions of "trans-science," see Wendy E. Wagner, *The Science Charade in Toxic Risk Regulation*, 95 COLUM. L. REV. 1613, 1618-28 (1995); David A. Wirth, *The Role of Science in the Uruguay Round and NAFTA Trade Disciplines*, 27 CORNELL INT'L L.J. 817, 832-40 (1994); Giandomenico Majone, *Science and Trans-Science in Standard Setting*, 9 SCI., TECH., & HUMAN VALUES, Winter 1984, at 15.

2. Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations, Apr. 15, 1994, 33 I.L.M. 1143 (1994) [hereinafter Final Act].

Measures (SPS Agreement),³ the question is to what extent the United States and other member states have ceded, or should cede, regulatory authority at this same interface to the World Trade Organization (WTO).⁴ The SPS Agreement sets forth the rights and obligations of members with respect to any measure taken by a member to protect the health or life of humans, animals, or plants ("sanitary or phytosanitary measures") that may, directly or indirectly, affect international trade.⁵ On the one hand, members have the sovereign right to take such measures to protect health and life within their territories; on the other hand, they may do so only if such measures are not inconsistent with the provisions of the SPS Agreement and, in particular, are not arbitrarily or unjustifiably discriminatory and do not constitute disguised restrictions on international trade.⁶ If disputes arise between WTO members concerning compliance with the SPS Agreement, they are to be resolved through the process set up by the Dispute Settlement Understanding (DSU).⁷ The DSU established a Dispute Settlement Body (DSB) of the WTO with authority to establish factfinding panels, adopt panel and Appellate Body reports, and otherwise administer the dispute settlement process of the WTO.⁸ Thus, WTO members agree that, in disputes over whether a member's domestic regulatory measures are inconsistent with the SPS Agreement, the DSB is the final arbiter.

The central strategy of the SPS Agreement is to use science to distinguish between those sanitary measures consistent with the Agreement and those in violation of the Agreement. In essence, a sanitary measure that adversely affects international trade is consistent with the SPS Agreement only to the extent it is necessary to protect health or life, and provided it "is based on scientific principles and is not maintained without sufficient scientific evidence."⁹ Such measures must also be "based on" risk assessment methods¹⁰ and must be undertaken pursuant to certain risk management

3. Agreement on the Application of Sanitary and Phytosanitary Measures, Apr. 15, 1994, annex 1A-4 to Final Act Embodying the Results of the Uruguay Round of Multilateral Trade Negotiations, in *Message From the President of the United States Transmitting the Uruguay Round Trade Agreements, Texts of Agreements, Implementing Bill, Statement of Administrative Action, and Required Supporting Statements*, H.R. Doc. No. 103-316, Vol. I, at 1381-95 [hereinafter SPS Agreement].

4. The WTO was established during the 1994 Uruguay Round of Multilateral Trade Negotiations under the Agreement Establishing the World Trade Organization, 33 I.L.M. 1144 (1994) [hereinafter WTO Agreement].

5. SPS Agreement, *supra* note 3, preamble, art. 1.1, annex A.1.

6. *Id.* arts. 2.1, 2.3.

7. See Uruguay Round Understanding on Rules and Procedures Governing the Settlement of Disputes, 33 I.L.M. 1226, arts. 1.1, 23, app. 1 (1994) [hereinafter DSU]; WTO Agreement, *supra* note 4, art. 4.3; SPS Agreement, *supra* note 3, art. 11.1. Unless otherwise indicated, references in this Article to a "member" refer to membership under the WTO Agreement, *supra* note 4, and the DSU, *supra*.

8. DSU, *supra* note 7, art. 2.1.

9. SPS Agreement, *supra* note 3, art. 2.2. Cf. Julie Cromer, *Sanitary and Phytosanitary Measures: What They Could Mean for Health and Safety Regulations Under GATT*, 36 HARV. INT'L L.J. 557, 568 (1995) (stating that the SPS Agreement's purpose is "to ensure that free trade continues regardless of technological differences or expectations as to what is 'healthy' or 'safe'").

10. See SPS Agreement, *supra* note 3, arts. 5.1-5.3.

objectives and constraints.¹¹ Additionally, members hope to promote a global harmonization of all such sanitary and phytosanitary measures, with a central harmonizing role to be played by standards, guidelines, and recommendations established by certain international institutions.¹²

The interpretation of nearly all of these important provisions of the SPS Agreement was at issue in a recent proceeding under the DSU. In response to complaints by the United States and Canada against the European Communities, a WTO factfinding panel (the "Hormones Panel") found that the European Communities violated the SPS Agreement by maintaining and implementing measures that prohibit the marketing of meat from farm animals treated with certain hormones for the purpose of increasing feed efficiency and the animal's rate of growth ("growth promotion" purposes).¹³ The primary European justification for these measures was that the hormones are carcinogenic¹⁴ and that growth promotional use adds to the risk already faced by consumers from background levels of hormones.¹⁵ The European Communities appealed the adverse reports of the WTO Hormones Panel, and the Appellate Body issued a single report on January 16, 1998.¹⁶ The DSB adopted the Appellate Body report, as well as the Panel reports as modified by the Appellate Body.¹⁷

The growth hormones dispute is not only a difficult case of first impression under the SPS Agreement, but it also involves legal issues about the interface between science and regulation. Moreover, it presents these issues in a critical international context.¹⁸ One way of posing these issues

11. See *id.* arts. 5.3-5.6.

12. *Id.* preamble, art. 3, annex A.3.

13. Report of the Panel, EC Measures Concerning Meat and Meat Products (Hormones) – Complaint by the United States, Aug. 18, 1997, WT/DS26/R/USA, available in Westlaw, WTO-DEC file, 1997 WL 569984 [hereinafter Hormones Panel Report (US)]. On April 25, 1996, the United States requested the DSB to establish a factfinding panel. Hormones Panel Report (US) ¶ 1.4. Report of the Panel, EC Measures Concerning Meat and Meat Products (Hormones) – Complaint by Canada, Aug. 18, 1997, WT/DS48/R/CAN, available in Westlaw, WTO-DEC file, 1997 WL 561677 [hereinafter Hormones Panel Report (CAN)]. On September 27, 1996, Canada requested the DSB to establish a panel. Hormones Panel Report (CAN) ¶ 1.3. Copies of both reports are available on the Internet at *World Trade Organization* (visited Apr. 21, 1998) <<http://www.wto.org>>.

The issues in the Canadian and U.S. proceedings were largely the same, and the two reports are highly parallel in content. The Canadian report is the longer of the two (257 single-spaced pages, compared to the U.S. report of 225 pages). I have, therefore, relied principally on the Canadian report for quotations, but have given citations to parallel paragraphs in the U.S. report when it seemed helpful to do so.

14. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶ 3.6.

15. See, e.g., *id.* ¶ 8.192.

16. Report of the Appellate Body, EC Measures Concerning Meat and Meat Products (Hormones), Jan. 16, 1998, adopted Feb. 13, 1998, WT/DS26/AB/R & WT/DS48/AB/R, available in Westlaw, WTO-DEC file, 1998 WL 25520 [hereinafter Appellate Body Report].

17. See *Overview of the State-of-play of WTO Disputes* (visited Apr. 21, 1998) <<http://www.wto.org/wto/dispute/bulletin.htm>>.

18. Wirth states that the "serious disagreement between the United States and the European Union over hormone-treated beef, now [in 1994] nearly a decade in duration, motivated much of [the SPS] text." Wirth, *supra* note 1, at 824.

is to ask whether the WTO should become, or must become, the "World Trans-science Organization," a global meta-regulator. In such a capacity, the WTO would resolve scientific issues such as carcinogenicity, adopt policies concerning the acceptable levels of risk or scientific uncertainty, or would make decisions about appropriate levels of health and safety. I argue in this Article that the WTO has no authority under the SPS Agreement to do any of these things. But this, in a sense, is not the difficult part of the argument. The difficult task is to identify the precise role of the WTO at the fact-policy-decisional interface between science and regulation, such that the WTO can legitimately and usefully settle trade disputes like the growth hormones dispute without becoming the World Trans-science Organization. This Article presents an affirmative theory of the WTO's proper role in such disputes.

In Part I of the Article, I analyze the relevant provisions of the SPS Agreement. In that context, I examine the distinction between risk assessment and risk management, as well as the concepts of scientific uncertainty and science policy. Part II sets forth my proposal for how the WTO should review the sanitary measures established by members. The focus of attention is the factfinding process in dispute settlement proceedings and, in particular, the substantive issues that should come before a panel, the evidence that a panel should consider, and the proper standard and burdens of proof. Finally, in Part III, I analyze the reports of the Hormones Panel and argue that the Panel made numerous and serious errors of law. I also argue that the Appellate Body, while it clarified a number of issues, also left many important questions unresolved. I conclude by suggesting the proper role of the WTO in factfinding about the regulatory programs of WTO members.

I. Obligations of Members Under the SPS Agreement

The SPS Agreement recognizes the sovereign right of every member to protect human, animal, or plant life and health within its territory,¹⁹ and addresses two distinct aspects of a risk-centered process for providing such protection. First, a member is expected to engage in *risk assessment*. The goals of risk assessment are to determine the adverse effects on human health that can be caused by exposure to a toxic agent, and to determine the potential for such exposure and adverse effects to occur.²⁰ Second, a member may undertake *risk management*: "the process of identifying, evaluating, selecting, and implementing actions to reduce risk."²¹ Risk man-

19. SPS Agreement, *supra* note 3, preamble, arts. 1.1, 2.1, annex A.1. Although the SPS Agreement applies to measures designed to protect the health or life of plants, animals, or humans, this Article will focus on the Agreement's protection of human life or health. The adoption of this shortened form is partly for rhetorical convenience, and partly because the WTO hormones dispute discussed here addressed almost exclusively the effects on humans. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶ 8.22.

20. SPS Agreement, *supra* note 3, arts. 5.1-5.3, annex A.4.

21. The Presidential/Congressional Commission on Risk Assessment and Risk Management, 1 FRAMEWORK FOR ENVIRONMENTAL HEALTH RISK MANAGEMENT 1 (1997). The National Research Council used the term "risk management":

agement involves the determination and application of an "acceptable level of risk" — that is, "the level of protection deemed appropriate by the member."²² The SPS Agreement sets forth certain requirements or disciplines for this process of assessing and managing risk. The Agreement also attempts to harmonize members' risk assessment and risk management decisions, insofar as doing so is consistent with a respect for sovereignty. The SPS Agreement therefore attempts to resolve the natural tension between the trade efficiencies of harmonization and the respect for sovereign protection of health, and it seeks to do so by appealing to "neutral" science. This part of the Article explores the roles of scientific uncertainty and science policy in risk assessment, and their relevance to the balance struck by the SPS Agreement between trade efficiencies and sovereign rights.

A. Scientific Uncertainty and Science Policy Within Risk Assessment

1. *The Nature of Risk Assessment*

Risk assessment involves a scientific determination of the relationships between causes and effects. It is generally agreed that total risk is a function of at least two factors: the toxicity of the agent and the predicted exposure to that agent.²³ Although exact terminology and lines of demarcation may differ, a risk assessment involves making scientific determinations in up to four distinct subcategories:²⁴

to describe the process of evaluating alternative regulatory actions and selecting among them. Risk management, which is carried out by regulatory agencies under various legislative mandates, is an agency decision-making process that entails consideration of political, social, economic, and engineering information with risk-related information to develop, analyze, and compare regulatory options and to select the appropriate regulatory response to a potential chronic health hazard. The selection process necessarily requires the use of value judgments on such issues as the acceptability of risk and the reasonableness of the costs of control.

NATIONAL RESEARCH COUNCIL, *RISK ASSESSMENT IN THE FEDERAL GOVERNMENT: MANAGING THE PROCESS* 18-19 (1983) [hereinafter NRC (1983)]. Taking action to reduce risk is the recognized right of all members. SPS Agreement, *supra* note 3, art. 2.1.

22. SPS Agreement, *supra* note 3, annex A.5.

23. See, e.g., the Presidential/Congressional Commission on Risk Assessment and Risk Management, *supra* note 21, at 1.

24. See generally NRC (1983), *supra* note 21; NATIONAL RESEARCH COUNCIL, *SCIENCE AND JUDGMENT IN RISK ASSESSMENT* (1994) [hereinafter NRC (1994)]; Commission Directive 93/67/EEC of 20 July 1993, L227 OFFIC. J. OF THE EUR. COMMUN. 9 (1993) [hereinafter Commission Directive] ("laying down the principles for assessment of risks to man and the environment of substances notified in accordance with Council Directive 67/548/EEC"); Report of the Joint FAO/WHO Expert Consultation, *Application of Risk Analysis to Food Standards Issues*, Geneva, Switzerland, Mar. 13-17, 1995 [hereinafter FAO/WHO Expert Consultation on Risk Analysis]; Codex Alimentarius Commission, Report of the Twenty-Second Session, Geneva, June 23-28, 1997, Agenda Item 6, Appendix II (adopting definitions of risk analysis terminology).

Not every risk assessment needs to encompass all four subcategories, NRC (1994), *supra*, at 27, but only "as appropriate to the circumstances," SPS Agreement, *supra* note 3, art. 5.1. The SPS Agreement does not specify the elements of a risk assessment in detail. See *id.* arts. 5.1-5.2, annex A.4. Although Article 5.1 of the Agreement directs members to take into account risk assessment techniques developed by relevant international organizations, the Hormones Panel found that, with regard to its proceeding, the

- *Hazard Identification* — identifying which environmental agents are capable of causing adverse health effects, as well as the types of adverse health effects those agents can cause;²⁵
- *Dose-Response Assessment*, or *Hazard Characterization* — evaluating the quantitative and qualitative aspects of the causal relationship between dose or level of exposure and the incidence or severity of the adverse health effects;²⁶
- *Exposure Assessment* — evaluating the probability, magnitude, duration, and timing of the doses that people might receive as the result of the various pathways of exposure to the agent;²⁷ and
- *Risk Characterization* — integrating information on hazard identification, dose-response assessment, and exposure assessment to develop a qualitative or quantitative estimate of the total risk of adverse health effects to a given population.²⁸

Each of these four elements will be examined in more detail below. Before doing so, however, it is necessary to develop the central concepts of scientific uncertainty and science policy.

The extent of the risk assessment required by the SPS Agreement in any particular case depends upon the circumstances,²⁹ including the nature and adequacy of the scientific information that is available. As long as any sanitary measure "is based on scientific principles"³⁰ and "is not maintained without sufficient scientific evidence,"³¹ and as long as the risk assessment on which that measure is based has taken into account the available information about toxicity and exposure,³² then the minimal requirements of the SPS Agreement are satisfied.

parties agreed that no particular techniques had to be taken into account. Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.106, 8.113.

25. NRC (1983), *supra* note 21, at 19, 20-23; NRC (1994), *supra* note 24, at 26, 57-60.

26. NRC (1983), *supra* note 21, at 19-20, 23-27; NRC (1994), *supra* note 24, at 26, 60-66; Commission Directive, *supra* note 24.

27. NRC (1983), *supra* note 21, at 20, 27-28; NRC (1994), *supra* note 24, at 26-27, 43-55; Commission Directive, *supra* note 24.

28. NRC (1983), *supra* note 21, at 20; NRC (1994), *supra* note 24, at 27, 68-78; Commission Directive, *supra* note 24.

29. See SPS Agreement, *supra* note 3, art. 5.1; *supra* note 24. Moreover, the member that relies on a risk assessment need not have conducted either that assessment or the scientific studies that are taken into account in making that assessment. See Appellate Body Report, *supra* note 16, ¶ 190. This is a matter of common sense and efficiency, as well as common practice. In many regulatory systems, the party seeking regulatory approval may be required to conduct the necessary studies and submit them for governmental review. See, e.g., 7 U.S.C. § 136a (1994) and 40 C.F.R. Part 158 (1997) (establishing data requirements for pesticide registration); 21 U.S.C. §§ 331, 342(a)(1)(C), 348 (1994) and 21 C.F.R. Part 171 (1997) (establishing data requirements for petitions proposing issuance of a regulation allowing use of food additive).

The SPS Agreement treats a member's particular "provisions on . . . methods of risk assessment" as sanitary measures to be evaluated under the Agreement. SPS Agreement, *supra* note 3, annex A.1. For the requirements on sanitary measures, see discussion *infra* Part I.B.

30. SPS Agreement, *supra* note 3, art. 2.2.

31. *Id.*

32. *Id.* art. 5.2.

Many, if not most, risk assessments today involve scientific uncertainty. Scientific uncertainty is due to a lack of knowledge, and therefore reflects the potential for error inherent in scientific information.³³ In even routine conclusions about causation, there are inherent uncertainties resulting from the choice of the particular variables used to gather data, the measurements taken, the samples drawn, the mathematical models used, and the causal factors and relations posited.³⁴ The U.S. Environmental Protection Agency (EPA) states that uncertainty in risk assessment can be classified into three major categories: measurement uncertainty, uncertainty associated with the use of scientific models (e.g., dose-response models), and data gaps.³⁵ Alternatively, a simple distinction can be made between uncertainty about which model to use and uncertainty about which values to use as inputs to those models.³⁶ Regardless of which taxonomic scheme we use for describing uncertainty, it is clear that scientific uncertainty pervades the empirical sciences and virtually all current estimates of risk.³⁷ There is an inherent potential for error associated with the validity or reliability of the available data, the predictive accuracy of the mathematical models used, and the explanations about the causal mechanisms at work.

In order to complete a risk assessment that adequately characterizes such uncertainty, risk assessors usually evaluate the scientific plausibility of alternative models or of alternative model inputs. Risk assessors may discount some accounts as not based on scientific principles or as inconsistent with established scientific theories. It is the nature of scientific uncertainty, however, that two or more alternative accounts will remain within the realm of plausibility.³⁸ Moreover, it seems impossible at the present time to provide a single test for judging whether a conclusion is "scientifically plausible." Roughly speaking, a causal account can be said to be scientifically plausible whenever it is supported by empirical data (as

33. See U.S. Environmental Protection Agency (EPA), *Memorandum: EPA Risk Characterization Program*, Mar. 21, 1995, available on the Internet at *Environmental Protection Agency* (last visited Apr. 21, 1998) <<http://www.epa.gov/ordntrnt/ORD/spc/rccover.htm>>; EPA, *Policy for Risk Characterization*, Mar. 1995, available on the Internet at *Environmental Protection Agency* (last visited Apr. 21, 1998) <<http://www.epa.gov/ordntrnt/ORD/spc/rcpolicy.htm>>; EPA, Science Policy Council, *Guidance for Risk Characterization*, Feb. 1995, available on the Internet at *Environmental Protection Agency* (last visited Apr. 21, 1998) <<http://www.epa.gov/ordntrnt/ORD/spc/rcguide.htm>>; NRC (1994), *supra* note 24, at 28-29, 161-62; FAO/WHO Expert Consultation on Risk Analysis, *supra* note 24, at 11-13, 24-29; Vern R. Walker, *The Siren Songs of Science: Toward a Taxonomy of Scientific Uncertainty for Decisionmakers*, 23 CONN. L. REV. 567, 570-71 (1991); NRC (1983), *supra* note 21, at 11-13.

34. Each of these types of uncertainty is discussed in Walker, *supra* note 33, at 574-618.

35. *Guidance for Risk Characterization*, *supra* note 33, at 8-9.

36. See NRC (1994), *supra* note 24, at 11-12, 162-66; FAO/WHO Expert Consultation on Risk Analysis, *supra* note 24, at 12, 25.

37. See *Guidance for Risk Characterization*, *supra* note 33, at 19-20; Report of a Joint FAO/WHO Consultation, *Risk Management and Food Safety*, Rome, Italy, January 27-31, 1997, at 8 [hereinafter FAO/WHO Consultation on Risk Management]; Walker, *supra* note 33, at 572-73, 625-26.

38. See, e.g., NRC (1983), *supra* note 21, at 28-37.

opposed to mere speculation or personal intuition) and by a line of reasoning (often including a model and theory), which together provide a rational basis for drawing a conclusion, even though reasonable scientists might disagree on whether that conclusion is the only inference that can be drawn validly from the data.³⁹ Thus, there can be several scientifically plausible conclusions or accounts, with wide disagreement among scientists as to which conclusions or accounts will ultimately prove to be correct.⁴⁰ As indirect evidence, therefore, the existence of good-faith disagreement among respected scientists is itself a good indication that the alternative accounts are scientifically plausible.

In the presence of such scientific uncertainty, risk assessment often proceeds by choosing from among the alternative accounts that remain scientifically plausible once all of the available scientific information has been considered. The following part of this Article will illustrate that such uncertainties are both numerous and inherent in scientific judgments under each of the four elements of risk assessment. Faced with the reality of limited knowledge, "assessors must [in many cases] choose among available data, models or assumptions in estimating risks."⁴¹ When science does not provide a definitive answer as to which data, models or assumptions should be used, EPA encourages risk assessors to follow guidance from the Agency's "science policy" in conducting risk assessments.⁴² An

39. This formulation is necessarily broad and programmatic, because details depend upon the particular area of research, the particular sciences involved, and the current state of scientific information. For a discussion of plausible accounts in the area of carcinogenicity, see *infra* Part I.A.2.

40. See NRC (1983), *supra* note 21, at 36.

41. *Guidance for Risk Characterization*, *supra* note 33, at 4-11.

42. "Science policies" are determinations about how risk assessors should proceed when they encounter uncertainties involving multiple plausible accounts. Because such policies usually specify which assumptions to use to bridge gaps in our scientific knowledge, they are sometimes called "inference guidelines" or "default assumptions." See NRC (1983), *supra* note 21, at 28-37; NRC (1994), *supra* note 24, at 27, 85-90. As EPA has stated:

The National Research Council, in its 1983 report . . . recognized that default assumptions are necessarily made in risk assessments where gaps exist in general knowledge or in available data for a particular agent. These default assumptions are inferences based on general scientific knowledge of the phenomena in question and are also matters of policy concerning the appropriate way to bridge uncertainties that concern potential risk to human health . . . from the agent under assessment.

Guidelines for Carcinogen Risk Assessment, 61 Fed. Reg. 17,960, 17,964 (proposed Apr. 23, 1996) [hereinafter Proposed EPA Carcinogen Guidelines]. See FAO/WHO Expert Consultation on Risk Analysis, *supra* note 24, at 11-13, 24-29 (Risk assessment "requires explicit recognition of uncertainties and, when appropriate, acknowledgement that alternative interpretations of the available data may be scientifically plausible."). Cf. Weinberg, *supra* note 1, at 215-16, 220 (stating that important contribution of scientists is "to say where science ends and trans-science begins").

EPA and other authorities have expressly recognized the role of science policy in conducting risk assessment. As EPA stated:

[w]ithin the Risk Assessment category there is a group that develops chemical-specific risk assessments by collecting, analyzing, and synthesizing scientific data to produce the hazard identification, dose-response, and exposure assessment portion of the risk assessment and to characterize risk. This group relies

example of a science policy is the presumption that a chemical agent that can cause cancer in laboratory animals can also cause cancer in humans. Numerous examples of such science policies are given in the following part of this Article. EPA has issued several risk assessment guidelines that identify the science policies (and resulting "inference guidelines" or "default assumptions") to be used by Agency risk assessors.⁴³ It is also explicit EPA policy that the risk characterization step of a risk assessment should include a full discussion of the scientific uncertainties underlying all of the steps of the risk assessment, and that the risk characterization should identify the science policies or guidelines used in making the assessment.⁴⁴

Science policies themselves are not justified on purely scientific grounds. In fact, the need for science policy arises precisely because our limited scientific knowledge permits multiple accounts that are scientifically plausible. The pragmatic goal of risk assessment is to characterize faithfully the current state of scientific knowledge, and to do so in a manner and in a time frame that is useful for risk management decision-making.⁴⁵ Therefore, while risk assessment should be as purely scientific and as free of policy as possible, the state of scientific knowledge in a given

in part on Agency risk assessment guidelines to address science policy issues and scientific uncertainties.

Guidance for Risk Characterization, *supra* note 33, at 1. The National Research Council, in its foundational 1983 report on risk assessment in the Federal Government, stated that: "[t]he choices encountered in risk assessment rest, to various degrees, on a mixture of scientific fact and consensus, on informed scientific judgment, and on policy determinations (the appropriate degree of conservatism)." NRC, *supra* note 21, at 36. In a subsequent report, the National Research Council stated that:

[l]ike the committee that produced the 1983 NRC report, we recognize that there is an inevitable interplay between risk assessment and risk management. As the 1983 report states (pp. 76, 81), "risk assessment must always include policy, as well as science," and "guidelines must include both scientific knowledge and policy judgments." Any choice of defaults, or the decision not to have defaults at all, therefore amounts to a policy decision.

NRC (1994), *supra* note 24, at 87.

43. For example, in the area of risk assessment for carcinogenicity, the EPA has published Guidelines for Carcinogen Risk Assessment, 51 Fed. Reg. 33,992 (1986), has proposed revised guidelines, and has stated that "[b]oth the 1986 guidelines and the current [1996] proposal contain inference guidance in the form of default inferences to bridge gaps in knowledge and data." Proposed EPA Carcinogen Guidelines, *supra* note 42, at 17,960. On the terminology of "inference guidelines" or "default assumptions," see *supra* note 42.

44. See *Guidance for Risk Characterization*, *supra* note 33. EPA stated that:

This section focuses on two requirements for full characterization of risk. . . . Second, [the risk characterization] should identify the important strengths and uncertainties in the assessment as part of a discussion of the confidence in the assessment.

. . .

A discussion of uncertainty requires comment on such issues as the quality and quantity of available data, gaps in the data base for specific chemicals, quality of the measured data, use of default assumptions, incomplete understanding of general biological phenomena, and scientific judgments or science policy positions that were employed to bridge information gaps.

Id. at 4-5.

45. See Memorandum: EPA Risk Characterization Program, *supra* note 33; Policy for Risk Characterization, *supra* note 33, at 1-3.

area is often so limited that risk assessors are forced to choose among alternative models or inputs by following rules that take into account considerations other than science.⁴⁶ These rules are science policies. For example, in the face of scientific uncertainty, it is common to select models or model inputs that analysts consider to be conservative in protecting human health, although the decision to protect human health is not a canon of pure science.⁴⁷ So while risk assessors who arrive at a risk characterization in the face of scientific uncertainty should be guided by science policies, those policies are not themselves scientific in nature and are not created by risk assessors alone. Rather, they are policies that reflect the broader goals of risk regulation, such as protecting human health.

Explicit science policies or inference guidelines allow risk assessments to remain "objective" by maintaining consistency and transparency in the face of scientific uncertainty, even though some risk management goals are used to provide guidance to risk assessors about how to proceed.⁴⁸ We are able to maintain a functional distinction between risk assessment and risk management precisely because the former is limited to drawing inferences supported by science alone, or by science combined with explicit science policies or inference guidelines.⁴⁹

EPA recognizes that while risk assessors should follow science policy guidelines when choosing among scientifically plausible alternatives and when reporting underlying uncertainties within the risk characterization,

46. See *supra* note 42. The central purpose of science policies is to bridge gaps in current scientific knowledge. Therefore, when there is an advance in scientific knowledge that fills a particular gap, the policy should be replaced by that knowledge. Until that time comes, however, science-policy choices made by risk assessors should always be choices from among scientifically plausible alternatives, and each alternative must have a reasonable scientific basis in order to be considered scientifically plausible.

47. See *Guidance for Risk Characterization*, *supra* note 33, at 1-10.

48. See *Memorandum: EPA Risk Characterization Program*, *supra* note 33, at 2. EPA stated that:

[B]ecause transparency in decisionmaking and clarity in communication will likely lead to more outside questioning of our assumptions and science policies, we must be more vigilant about ensuring that our core assumptions and science policies are consistent and comparable across programs, well grounded in science, and that they fall within a "zone of reasonableness."

Id.

As stated by a Joint FAO/WHO Consultation: "Risk assessment policy setting is a risk management responsibility, which should be carried out in full collaboration with risk assessors, and which serves to protect the scientific integrity of the risk assessment. The guidelines should be documented so as to ensure consistency and transparency." FAO/WHO Consultation on Risk Management, *supra* note 37, at 5.

49. The alternatives from which science-policy choices are made should qualify as candidates solely on the basis of scientific principles, even though the choices themselves are guided by risk management objectives. Thus, risk management and risk assessment necessarily interact, see NRC (1994), *supra* note 24, at 87, in what should be transparent and precise ways, while they also maintain distinct roles. See *Guidance for Risk Characterization*, *supra* note 33, at 2 ("Risk assessors and risk managers should be sensitive to distinctions between risk assessment and risk management. . . . Where responsibilities differ, it is important that participants confine themselves to tasks in their areas of responsibility and not inadvertently obscure differences between risk assessment and risk management.").

they should not take on the role properly reserved for risk managers.⁵⁰ For example, in selecting, evaluating, and presenting scientific information, risk assessors should not consider issues such as cost or feasibility or how the scientific analysis might influence the regulatory decision.⁵¹ Additionally, risk assessors are not to "make decisions on the acceptability of any risk level for protecting public health or selecting procedures for reducing risks."⁵² EPA states that: "[m]atters such as risk assessment priorities, degree of conservatism, and acceptability of particular risk levels are reserved for decision-makers who are charged with making decisions regarding protection of public health."⁵³ Such management decision-making is reserved for legislatures or governmental officials charged with protecting public health. It falls to risk management to balance competing societal goals (including costs and benefits) and to take science, scientific uncertainty, and risk characterization into account in setting generic science policies and in arriving at specific decisions about risk acceptance or reduction.⁵⁴

According to the United States, the SPS Agreement "recognizes the fact that scientific certainty is rare and many scientific determinations require judgments between differing scientific views. The [SPS] Agreement preserves the ability of governments to make such judgments."⁵⁵ All scientifically plausible alternatives have, by definition, a reasonable scientific basis and, to use the phrasing of the SPS Agreement, are "based on scientific principles," have "sufficient scientific evidence," and "take into account available scientific evidence."⁵⁶ On the other hand, any of several scientifically plausible alternatives might eventually prove to be the correct account. Therefore, the SPS Agreement requires a risk determination based on scientifically plausible models and model inputs, but does not pretend that scientific evidence can always, or even often, prove which plausible alternative is the correct one.

50. See *Guidance for Risk Characterization*, *supra* note 33, at 2-3; cf. FAO/WHO Consultation on Risk Management, *supra* note 37, at 5-6, 15 (recommending that determination of "risk assessment policy" should be "a specific component of risk management," and that a "functional separation" should be maintained between risk management and risk assessment).

51. *Guidance for Risk Characterization*, *supra* note 33, at 2.

52. *Id.*

53. *Id.* at 3.

54. EPA stated that:

[r]isk managers, as a separate category [from risk assessors], *integrate* the risk characterization with other considerations specified in applicable statutes to make and justify regulatory decisions.

...

For example, a regulatory decision on the use of a particular pesticide considers not only the risk level to affected populations, but also the agricultural benefits of its use that may be important for the nation's food supply.

Id. at 2, 3.

55. The Uruguay Round Agreements Act, Statement of Administrative Action, Agreement on the Application of Sanitary and Phytosanitary Measures § 6.b, *reprinted in* 1994 U.S.C.A.N. 4102, 4105.

56. See SPS Agreement, *supra* note 3, arts. 2.2, 5.2.

Setting science policy, which specifies how to deal with scientific uncertainty within risk assessments, is itself a function of risk management. The SPS Agreement leaves room for members to choose different science policies. A given member, therefore, might disagree with the risk conclusions drawn by a study's scientific author, by the authors of a risk assessment report based on the study, or by another member considering the same data, and the SPS Agreement would not be violated so long as these disagreements are rooted in scientific principles. Whenever scientific uncertainty is present in risk assessment, each member should be entitled to choose between scientifically plausible options and should be able to follow its own science policies, which reflect in turn that member's management policies. The only alternative would be to require members to adopt a set of WTO science policies.⁵⁷

2. The Elements of Risk Assessment

Science policy determinations pervade risk assessment and are essential to completing most risk assessments. This point, which must be fully appreciated in order to interpret the SPS Agreement in a workable way, can be illustrated by a brief survey of EPA's position on this issue in the context of agency risk assessments for carcinogens.⁵⁸

a. Hazard Identification

The determination that a chemical agent is carcinogenic to humans usually involves many choices among scientifically plausible options. An important example is the appropriate weight to be placed on carcinogenicity data from studies of laboratory animals. EPA recognizes the following major default assumptions in its inferences from animal data to conclusions about the carcinogenicity of an agent in humans: that positive effects in animal cancer studies indicate carcinogenic potential in humans; that effects seen in animals at the highest dose tested (the maximum tolerated dose) are appropriate to use in carcinogenicity assessment; that concordance among target organs of carcinogenicity is not a prerequisite; that benign tumors observed in the animal studies should be included in assessing animal tumor incidence if the tumors have the capacity to progress to malignancies; that there is a similarity of the basic pathways of metabolism that are relevant to species-to-species extrapolation of cancer hazard; and that a human dose that is equivalent to an animal dose can be

57. Imagine, for example, that WTO dispute settlement panels were to rank competing but scientifically plausible alternatives on a scale ranging from "merely plausible" to "highly plausible." A panel finding that certain measures were based on "merely plausible" theories would find those measures to be inconsistent with the SPS Agreement. It would be hard to find a source of legal authority for such an undertaking. Moreover, there is no scientific consensus on a general method for performing such ranking, and it is currently implausible that such ranking could be scientifically demonstrated to be correct. If WTO panels were to begin finding some scientifically plausible options to be unacceptable bases under the SPS Agreement, then this would be an example of imposing a WTO science policy upon members.

58. See generally EPA Guidelines for Carcinogen Risk Assessment, *supra* note 43; Proposed EPA Carcinogen Guidelines, *supra* note 42.

estimated by using a scaling factor based on body weight (in the case of oral exposure), default estimates of lung deposition and of internal dose (for inhalation exposure), or internal dose (for a route-to-route of exposure extrapolation).⁵⁹ Unless such default rules were used to bridge large gaps in fundamental knowledge or smaller gaps in agent-specific knowledge, hazard identification and risk assessment could not be completed.⁶⁰ On occasion, such default assumptions have been incorporated into statutes, as in the case of the Delaney Clause governing approval of food additives under the U.S. Federal Food, Drug, and Cosmetic Act, which treats as human carcinogens any food additives that can cause cancer in laboratory animals.⁶¹

b. Dose-Response Assessment

If a chemical agent is capable of causing an adverse health effect in humans at some dose, then it is important to determine the conditions under which that adverse effect can occur, and what the relevant dose or level of exposure is. For example, EPA recognizes the following major default assumptions in its inferences from animal data at high doses to conclusions about the expected carcinogenic response of humans at low doses: a biologically

59. Proposed EPA Carcinogen Guidelines, *supra* note 42, at 17,967-68.

60. See *id.* at 17,964-66 ("[T]he major default assumptions commonly employed in cancer risk assessment and adopted in these guidelines" are "predominantly inferences necessary to use data observed under empirical conditions to estimate events and outcomes under environmental conditions.").

61. 21 U.S.C.A. § 348(c)(3)(A) (West Supp. 1998). The statute provides that: "no additive shall be deemed to be safe if it is found to induce cancer when ingested by man or animal" (emphasis added). As the U.S. Government stated:

The determination that a particular substance poses a risk of cancer is a scientific determination, based on an evaluation of the potential for a substance to induce cancer. Based on scientific principles, the United States has determined that if a substance induces cancer in animals, it poses some risk of human carcinogenesis.

Statement of Administrative Action, *supra* note 54, § 9.

See Jeffery Atik, *Science and International Regulatory Convergence*, 17 NW. J. INT'L L. & BUS. 736, 744-45 (1996-97) (stating that SPS Agreement grants "nearly complete discretion to legislatures to enact SPS measures where risk assessment has been performed" and that a risk assessment might prove "adequate . . . for purposes of staving off a trade challenge" if the assessment takes notice that "a substance has been shown, to any degree, to cause cancer in laboratory animals"); Zane O. Gresham & Thomas A. Bloomfield, *Rhetoric or Reality: The Impact of the Uruguay Round Agreement on Federal and State Environmental Laws*, 35 SANTA CLARA L. REV. 1143, 1148-50 (1995) (stating that a decision to adopt "zero-risk tolerance," such as the U.S. Delaney Clause, is not subject to challenge under the SPS Agreement and that U.S. regulations "are based upon a political choice of what level of risk is appropriate and a scientific determination regarding whether the risk exists"); Robert M. Millimet, *The Impact of the Uruguay Round and the New Agreement on Sanitary and Phytosanitary Measures: An Analysis of the U.S. Ban on DDT*, 5 TRANSNAT'L L. & CONTEMP. PROBS. 449, 473 (1995) (arguing that the Delaney Clause is sufficiently premised on science to satisfy the SPS Agreement because it was enacted after scientific testimony that the then-current techniques were incapable of determining a safe level of carcinogens, and because the clause is triggered only by positive data showing carcinogenicity in animals). But see Jeffrey L. Dunoff, *Institutional Misfits: The GATT, the ICJ and Trade-Environment Disputes*, 15 MICH. J. INT'L L. 1043, 1076-77 (1994) (stating that Delaney Clauses rest "on the political — rather than scientific — judgment that no risk posed by carcinogens in the food supply is acceptable").

based or case-specific model is preferred for extrapolating below the range of observed data, but a curve-fitting model is to be used in the absence of such models; if a curve-fitting model is used and evidence on the mode of action is not sufficient to infer a threshold of action, then the default approach is to assume that carcinogenic response is linearly related to dose from the lowest observed data down to zero dose/zero response; and, if a curve-fitting model is used but there is sufficient evidence to infer a threshold of action, then the margin of exposure is calculated by dividing the low-dose-response point within the range of observation by the environmental exposure of interest.⁶² Without such assumptions — whether made in accordance with formally adopted science policies or ad-hoc by individual scientists — dose-response assessments for carcinogenicity could not be completed.

An especially critical regulatory task is to determine the shape of the dose-response curve at the low doses typical of many human exposures. For non-carcinogenic effects, regulators try to determine whether there is a "no-effect threshold" specific to the agent and the adverse effect — that is, a dose or exposure level below which the adverse effect does not occur and above which it can occur.⁶³ The question is whether there is a "safe" level of exposure below which there is no risk of adverse effect. If the agent is carcinogenic, however, there are often plausible scientific grounds for acting as though no threshold exists or for inferring that, if it exists, it is very low and cannot be identified reliably.⁶⁴ Especially when the causal mechanism for cancer is unclear or the agent is genotoxic, the traditional regulatory approach has been to treat carcinogens as not having a no-effect threshold.⁶⁵ Such carcinogens are treated as though any level of exposure poses some positive risk.

In extrapolating from high-dose effects to low-dose effects, from laboratory animals to humans, and from healthy individuals to more sensitive individuals, many choices have to be made among scientifically plausible options. Uncertainties arise repeatedly along the chains of inference, and science policies are adopted to guide the choices of scientists performing risk assessment. The numerous examples given above were adopted by EPA as a matter of Agency policy. Sometimes those choices are made by legislative institutions. For example, one Congressional science policy is

62. Proposed EPA Carcinogen Guidelines, *supra* note 42, at 17,968-69, 17,992-96. The EPA considers the choice of the appropriate margin of exposure to employ in regulation to be a question of risk management, not risk assessment. *Id.* at 17,969, 17,993. A typical and traditional margin of exposure might be 100 — a factor of 10 to make a conservative allowance for extrapolating from animals to humans, and an additional factor of 10 to allow for variability in response among human beings. *Id.* As EPA points out, these "10-fold factors are moderately conservative, traditional ones used for decades in the assessment of toxicological effects." *Id.* at 17,969. See NRC (1994), *supra* note 24, at 30-31.

63. NRC (1994), *supra* note 24, at 28-31.

64. See, e.g., Proposed EPA Carcinogen Guidelines, *supra* note 42, at 17,968-69, 17,992-93; NRC (1994), *supra* note 24, at 29-31, 65-66, 85-90.

65. See *infra* note 66; Junius C. McElveen, Jr., *Risk Assessment in the Federal Government: Trying to Understand the Process*, 5 TUL. ENVTL. L.J. 45, 62-67 (1991).

the Delaney Clause that prohibits the use of food additives that cause cancer in laboratory animals.⁶⁶ Whether those choices are made in statutes or regulations, or whether they are made in the form of generic guidelines for all decisions or on a purely ad-hoc and case-specific basis, they are essential to completing dose-response assessment.

c. Exposure Assessment

As noted above, risk is a function not only of the toxicity of the agent, but also of the likelihood that a dose will be received and the likely magnitude of that dose. The SPS Agreement acknowledges this need to determine exposure by requiring members to take into account "relevant processes and production methods" for products that might contain toxins, as well as "relevant inspection, sampling and testing methods" for those products.⁶⁷ In predicting the magnitude, duration, and frequency of exposure in actual populations, agencies such as EPA often have to make assumptions to bridge gaps in available input data.⁶⁸ They also have to adopt models for predicting the environmental fate of chemicals in air, water, and soil.⁶⁹ Science policy, therefore, can play as essential a role in exposure assessment as it does in toxicity assessment.

d. Risk Characterization

The objectives in risk characterization are to summarize the primary conclusions about hazard, dose-response, and exposure, to integrate those conclusions into an assessment of risk, and to present the integrated risk information in a form useful to decision-makers and risk managers.⁷⁰ EPA has been careful to emphasize that risk characterizations should also describe "the constraints of available data and the state of knowledge, significant scientific issues, and significant science and science policy choices that were made when alternative interpretations of data existed."⁷¹ Unless

66. 21 U.S.C.A. § 348(c)(3)(A) (West Supp. 1998). The U.S. Government stated that: "Based on scientific principles, the United States has determined that if a substance induces cancer in animals, it poses some risk of human carcinogenesis. And since the level of protection under Delaney requires that there be zero risk of carcinogenesis, the United States prohibits the substance." Statement of Administrative Action, *supra* note 55, § 9. As the National Research Council stated, the Delaney Clause is "[p]erhaps the earliest legislative acknowledgement of the possibility that chemical carcinogens might act in the same way" as radiation-induced cancer — that is, through nonthreshold mechanisms, such that "exposure to even one molecule of a carcinogen is associated with a small but non-zero increased risk of tumor induction." NRC (1994), *supra* note 24, at 31.

67. SPS Agreement, *supra* note 3, art. 5.2. As the Appellate Body stated:

[T]he risk that is to be evaluated in a risk assessment under [SPS] Article 5.1 is not only risk ascertainable in a science laboratory operating under strictly controlled conditions, but also risk in human societies as they actually exist, in other words, the actual potential for adverse effects on human health in the real world where people live and work and die.

Appellate Body Report, *supra* note 16, ¶ 187.

68. See EPA, Guidelines for Exposure Assessment, 57 Fed. Reg. 22,888, 22,909-17, 22,926-29 (1992).

69. *Id.*

70. See Policy for Risk Characterization, *supra* note 33, at 3-4.

71. Proposed EPA Carcinogen Guidelines, *supra* note 42, at 17,999-18,000.

a risk characterization acknowledges that "[r]isk assessment is an iterative process" and that "[d]efault assumptions are used at every stage because no database is ever complete," then risk assessment as a process is less likely to achieve its goals of "transparency in environmental decisionmaking, clarity in communication, consistency in core assumptions and science policies from case to case, and reasonableness."⁷² Every risk assessment of a potentially carcinogenic agent contains a great number of scientific uncertainties and default assumptions. These aspects would remain invisible to risk managers unless brought to the surface and explained. Risk assessment cannot remain true to its scientific ideals unless it also discloses the limitations of the science it contains. As the Greek founders of Western science emphasized, an essential aspect of true scientific knowledge is knowing what it is that we do not know.⁷³

B. Risk Management

Risk management is "the process of identifying, evaluating, selecting, and implementing actions to reduce risk to human health."⁷⁴ Risk managers decide what, if anything, to do about risk, by employing: (1) the results of risk assessment; (2) scientific information about the costs, benefits, and causal consequences of management choices; and (3) value judgments about societal goals and objectives.⁷⁵ The resulting risk management decisions balance such competing societal goals as taking a conservative approach to protecting human health and maximizing the net benefits of various forms of regulation. The difficulties of striking an appropriate balance are compounded by considerations of technological, political, and administrative feasibility, as well as by the inherent uncertainties in risk assessment. Such risk management decisions should be made by those in government who are charged with protecting public health,⁷⁶ including

72. See *id.* at 17,999.

73. In Plato's *Apology*, Socrates states:

[s]o when I went away, I thought to myself, "I am wiser than this man: neither of us knows anything that is really worth knowing, but he thinks that he has knowledge when he has not, while I, having no knowledge, do not think that I have. I seem, at any rate, to be a little wiser than he is on this point: I do not think that I know what I do not know."

PLATO, *EUTHYPHRO*, *APOLOGY*, *CRITO* 26 (F.J. Church trans., Bobbs Merrill Educational Publishing, 2d ed. 1956). Science policies should be formally adopted and risk assessment scientists should be required to disclose and explain inherent scientific uncertainties. In this way, those who make decisions based on a particular risk assessment will understand the limits of the underlying scientific knowledge. Absent such disclosure and explanation, decision-makers will not be able to distinguish guesswork from well-supported findings. See *supra* notes 42, 48.

74. The Presidential/Congressional Commission on Risk Assessment and Risk Management, *supra* note 21, at 1. See generally NRC (1983), *supra* note 21, at 18-19; FAO/WHO Consultation on Risk Management, *supra* note 37.

75. As the National Research Council stated, risk management "entails consideration of political, social, economic, and engineering information with risk-related information to develop, analyze, and compare regulatory options and to select the appropriate regulatory response." NRC (1983), *supra* note 21, at 19.

76. See EPA, *Guidance for Risk Characterization*, *supra* note 33, at 3 ("Matters such as risk assessment priorities, degree of conservatism, and acceptability of particular risk

national legislative bodies and administrative agencies.⁷⁷

The SPS Agreement recognizes at least three types of actions that a member may take to manage risks, and sets certain minimal requirements or disciplines for each. These three types of actions are:

- (1) Selecting the level of protection deemed appropriate by the member;
- (2) Establishing sanitary measures to achieve that level of protection; and
- (3) Accepting measures established by other members as being equivalent to its own.

Each of these functions will be discussed in turn.

1. *Selecting a Level of Protection*

The SPS Agreement defines the "appropriate level of protection" as that level of protection "deemed appropriate by the member" to protect human life or health within its territory.⁷⁸ Selecting the appropriate level of protection is an act of sovereignty, and the Agreement does not require the members to select any particular level.⁷⁹

The Agreement has only two provisions bearing on the selection of a

levels are reserved for decision-makers who are charged with making decisions regarding protection of public health.").

77. The "precautionary principle" employed in international environmental law is fundamentally a policy of risk management. A typical articulation of the precautionary principle is:

In order to protect the environment, the precautionary approach shall be widely applied by states according to their capabilities. 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.

Rio Declaration on Environment and Development, June 14, 1992, 31 I.L.M. 874, 879. For a survey of the various treaty articulations of the precautionary principle, see James E. Hickey, Jr. & Vern R. Walker, *Refining the Precautionary Principle in International Environmental Law*, 14 VA. ENVTL. L.J. 423, 431-38 (1995). The principle reflects the same precautionary approach so familiar from domestic environmental programs. As a policy of risk management, the precautionary principle is conservative in nature and strongly favors the protection of human health and the environment whenever such a goal must be balanced against considerations such as technological feasibility, economic efficiency, and administrative feasibility. See, e.g., *id.* at 425, 448-50 & n.102. For a brief discussion of the precautionary principle in the context of the SPS Agreement, see Wirth, *supra* note 1, at 838-40. But the conservative and health-protective approach of the precautionary principle can also function derivatively as a management goal in adopting science policies. When science policy is being established for risk assessment itself, the precautionary principle can guide risk assessors when they must make choices among scientifically plausible alternatives in the face of scientific uncertainty.

For a discussion of the relevance of the precautionary principle to the SPS Agreement, see Appellate Body Report, *supra* note 16, ¶¶ 120-25. As the Appellate Body stated: a panel charged with determining, for instance, whether "sufficient scientific evidence" exists to warrant the maintenance by a Member of a particular SPS measure may, of course, and should, bear in mind that responsible, representative governments commonly act from perspectives of prudence and precaution where risks of irreversible, e.g. life-terminating, damage to human health are concerned.

Id. ¶ 124.

78. SPS Agreement, *supra* note 3, annex A.5.

79. *Id.* preamble.

level of protection. First, the Agreement "encourages"⁸⁰ each member to "take into account the objective of minimizing negative trade effects."⁸¹ This provision is merely hortatory. The same objective is achieved by requiring that any adopted sanitary measures must not be "more trade-restrictive" than reasonably necessary.⁸² Thus, with respect to minimizing negative trade effects, the SPS Agreement places restrictions on the means chosen to achieve a selected level of protection, rather than on the process of selecting that level.

Second, the Agreement requires a member to be internally "consistent" in selecting levels of protection, and to avoid any "arbitrary or unjustifiable distinctions" in its levels of protection that would "result in discrimination or a disguised restriction on international trade."⁸³ This consistency requirement depends, therefore, on the meaning of "arbitrary or unjustifiable." The SPS Agreement recognizes that these concepts are difficult to define, and establishes a standing Committee to work with members to develop guidelines for implementing this provision.⁸⁴ The difficulty of determining consistency among levels of protection involving different substances, different adverse effects, and different products is indicated by the example that appears in the Agreement: that "the exceptional character of human health risks to which people voluntarily expose themselves" is a relevant factor in determining consistency.⁸⁵ The United States interprets this provision to mean that a member could "legitimately establish a high level of protection for pesticide exposure even if it sets a lower level of protection from the risks of cigarette smoking."⁸⁶ This example displays an appropriate recognition of the growing empirical data showing that factors such as the "voluntariness" of risk are important in predicting public reaction to that risk.⁸⁷ The difficulty in determining consistency derives from the sovereign and value-driven nature of the selection process. We are often hard pressed to explain why an individual takes some risks rather than others, or values some objectives more than others — and hard pressed to justify calling some sets of decisions "consistent" and other sets "inconsistent."⁸⁸ Evaluation of consistency at the individual level is difficult, but the difficulty is compounded when we evaluate such decisions at the societal level. Perhaps with the passage of time and the

80. Statement of Administrative Action, *supra* note 55, § 9.

81. SPS Agreement, *supra* note 3, art. 5.4.

82. See *infra* Part I.B.2.

83. SPS Agreement, *supra* note 3, art. 5.5.

84. *Id.*

85. *Id.*

86. Statement of Administrative Action, *supra* note 55, § 9.

87. See, e.g., Paul Slovic, *Perception of Risk*, 236 SCIENCE 280, 281-83 (1987).

88. For the growing literature on comparative risk assessment and countervailing risks, see, for example, RISK VERSUS RISK: TRADEOFFS IN PROTECTING HEALTH AND THE ENVIRONMENT (John D. Graham & Jonathan Baert Wiener eds., 1995); David A. Wirth & Ellen K. Silbergeld, *Risky Reform*, 95 COLUM. L. REV. 1857 (1995) (book review) (discussing conceptual difficulties with comparative risk assessment, especially in context of scientific uncertainty). Comparative risk assessments pose the same need for science policy as do single risk assessments, but encounter additional conceptual problems.

decision of numerous cases under the DSU, standards will evolve for deciding when different levels of protection are "arbitrary or unjustifiable." In the near term, however, such determinations should perhaps be made only in the most blatant and unexplainable cases.⁸⁹

One factor that seems clearly relevant to selecting a level of protection is the benefit to be gained from undertaking the associated risk.⁹⁰ Determining the "acceptable level of risk" is at least a function of prudential considerations under all of the circumstances, some of which we may be able to quantify in a cost-benefit analysis. In U.S. law, for example, a Delaney Clause prohibits the use of food additives found to induce cancer in animals,⁹¹ but EPA is permitted to consider the benefits to be obtained from the use of a particular carcinogenic pesticide in order to justify allowing a residue of that pesticide in food.⁹² The levels of protection are different, but that difference might be justifiable in terms of benefits.

The precise logical relationship of science policy to the selection of an appropriate level of protection is difficult to describe. However, it does seem reasonably clear that adopting policies to bridge scientific uncertainties is closely related to this aspect of risk management. Adopting a conservative approach to risk assessment, for example, should be seen as merely one aspect of selecting a high level of protection: being conservative in estimating risk is one aspect of selecting a highly protective management policy.⁹³ Moreover, the requirement of consistency between different

89. One of the intellectually elusive aspects of formulating the precautionary principle of international environmental law as a foundation of risk management is the difficulty of prescribing, generically and in advance, precisely how the protection of human health or the environment should be balanced against such considerations as economic efficiency and political feasibility. The multifactorial nature of such decisions renders it unlikely that in the near future there will be agreement on a single, adequate, and informative formulation of the principle that is enforceable as international law. See Hickey & Walker, *supra* note 77, at 424-26.

90. See, e.g., 7 U.S.C.A. § 136(bb) (West Supp. 1998) (defining "unreasonable adverse effects" for purposes of the U.S. pesticide statute).

91. 21 U.S.C.A. § 348(c)(3)(A) (West Supp. 1998). As the U.S. has stated: "The Delaney clauses, in the first instance, establish a level of protection. They reflect a decision by the Congress that there should be no risk of cancer to humans from the substances those clauses cover. That decision is fully protected under the [SPS] Agreement." Statement of Administrative Action, *supra* note 55, § 9.

92. See 7 U.S.C.A. §§ 136(bb), 136a(c)(5)(D) (West Supp. 1998); 21 U.S.C.A. § 346a(b)(2)(B)(iii)(II) (West Supp. 1998).

93. Alternatively, adopting a conservative science policy could be seen as a *means* of achieving a selected level of protection. However, it would be inappropriate to conceptualize science policies as "sanitary measures." First, doing so would subject science policies themselves to the disciplines of the SPS Agreement reserved for sanitary measures. Among those disciplines are the requirements that measures be "based on" a risk assessment, be "based on" scientific principles, and be "not maintained without sufficient scientific evidence." SPS Agreement, *supra* note 3, arts. 5.1, 2.2. But such requirements would be senseless and circular if applied to science policies. Second, although the SPS Agreement lists "methods of risk assessment" as an example of a sanitary measure, SPS Agreement, *supra* note 3, annex A.1, the word "methods" in that context should be read as referring to procedures, such as testing protocols, and not to the inference rules for interpreting the results of those procedures. Therefore, although science policies are clearly products of risk management (not merely of risk assessment), they

levels of protection is also a very sensible requirement for science policies. The adoption of science policies to be applied generically in risk assessment, wherever the relevant type of scientific uncertainty arises, would also make selecting a level of protection more transparent, and could assist in the development of criteria for when differences are "arbitrary or unjustifiable." In other words, the decision by members to handle scientific uncertainties through formally adopted science policies should assist in the evolution of SPS standards for selecting consistent levels of protection. For these reasons, we should regard the adoption of science policies as one aspect of selecting a level of protection. To the extent that scientific uncertainties exist and science policies are in play, the same deference that is due to a member's selection of a level of protection should be given to the member's selection of science policies to guide risk assessment.

2. *Establishing Measures to Achieve the Selected Level of Protection*

Although sovereign decisions to select an appropriate level of protection are largely unreviewable under the SPS Agreement, the Agreement places certain requirements on sanitary measures adopted as means for achieving a selected level of protection. First, a member should adopt "trade-restrictive" measures "only to the extent necessary" to achieve that member's selected level of protection.⁹⁴ This requirement has two aspects: effectiveness and efficiency. The only acceptable justification for a measure that is trade-restrictive is that it is reasonably effective in bringing about the targeted level of protection.⁹⁵ If a trade-restrictive measure does not in fact help to protect, then it cannot be justified as a means of providing protection. In addition, any sanitary measure should be reasonably efficient in bringing about that protection, in the sense of minimizing collateral effects on international trade.⁹⁶ A member may take effective measures to achieve its level of protection, so long as there are no alternative means to achieve the same level of protection that are "significantly less restrictive to trade."⁹⁷ The modifier "significantly less restrictive" suggests that *de minimis* differences between the effects of alternative measures are of no consequence to the SPS Agreement. A member is entitled to deference in its choice of means within some reasonable range of effects.

Second, a member should ensure that any measure is "based on" a risk assessment appropriate to the circumstances, is "based on" scientific principles, and "is not maintained without sufficient scientific evidence."⁹⁸ This second set of requirements is designed to ensure that a measure has a

are not themselves sanitary measures. Rather, they result from the sovereign function of selecting an appropriate level of protection.

94. SPS Agreement, *supra* note 3, arts. 2.2, 5.6.

95. A measure that does not contribute to bringing about an objective could hardly be considered necessary to that end. See WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY OF THE ENGLISH LANGUAGE 1510-11 (Unabridged) (1986) (defining "necessary" as meaning "cannot be done without" and "absolutely required").

96. See SPS Agreement, *supra* note 3, art. 5.6.

97. *Id.* art. 5.6 n.3.

98. *Id.* arts. 5.1, 2.2.

legitimate risk-reduction or risk-avoidance goal, and that the measure is grounded on the type of risk assessment foundation discussed above. It also reinforces the requirement of a reasonable cause-and-effect relationship between the sanitary measure chosen as a means and the societal objective defined in terms of a selected level of protection.

When these two requirements are considered together, we find another point of interaction between risk assessment and risk management, as well as another place to expect scientific uncertainty and the use of science policy to bridge gaps in knowledge. Under the Agreement, alternative measures are to be evaluated in part on their effectiveness in bringing about risk reduction.⁹⁹ For toxic agents, this is usually achieved by reducing or avoiding exposure to the agent, although occasionally toxicity itself is reducible.¹⁰⁰ In effect, in any comparison of alternative measures, we expect a risk assessment to be performed on each alternative and the resulting risk characterizations to be compared. In this process, the scientific uncertainties inherent in the risk assessments will be dealt with by science policies, which in turn are guided by the management policies at work in selecting an appropriate level of protection. The issue of how to deal with scientific uncertainties when two or more risk characterizations are to be compared may well give rise to additional science policies.

3. *Accepting Alternative Measures as Equivalent*

The SPS Agreement anticipates that different members might establish different sanitary measures. The Agreement respects the sovereign right of an importing member to select its own level of protection. It also defers to that member's choice of sanitary measures to achieve that level of protection, so long as the conditions stated in the previous part are met. The Agreement requires an importing member to accept the measures of another member as "equivalent" to its own measures — but only if "the exporting member objectively demonstrates to the importing member that its measures achieve the importing member's appropriate level of . . . protection."¹⁰¹ Thus, the Agreement is consistent in its deference to the importing member's cause-and-effect determinations when they are made in the context of the member's protection of the health of its own population. The SPS Agreement requires this deference so long as the importing member's determinations are based on scientifically plausible accounts of cause-and-effect.

C. Harmonizing Sanitary Measures Using International Standards

There is a difficult tension within the SPS Agreement between a respect for the sovereign right of members to protect health and a desire to increase trade efficiency through global harmonization of SPS measures. Moreover, the existence of scientific uncertainty and the resulting need for science

99. See *id.* arts. 2.2, 5.1, 5.6.

100. Legislation or regulation that prohibits the use of a toxic agent reduces risk by eliminating certain pathways of exposure. See, e.g., the Delaney Clause prohibition on carcinogenic food additives, *supra* note 66 and accompanying text; see *supra* Part I.A.2.c.

101. SPS Agreement, *supra* note 3, art. 4.1 (emphasis added).

policy complicate the Agreement's attempt to rely upon science as a neutral mediating principle. This part of the Article briefly examines the strategy under the SPS Agreement to use international standards, where they exist, to encourage but not impose harmonization.

The SPS Agreement does not mandate harmonization of all members' levels of protection. In fact, as noted above, decisions to select levels of protection are sovereign decisions that are accorded substantial deference under the Agreement.¹⁰² As discussed above in Part I.B, explicit SPS requirements are imposed principally on the establishment of sanitary measures as means of achieving the selected levels of protection. However, the Agreement promotes "international standards, guidelines or recommendations" as setting a *minimum* level of protection, while allowing members to set and maintain higher levels of protection.¹⁰³ As the United States has stated, the SPS Agreement "does not require 'downward harmonization' to less stringent [SPS] measures," and "a government is not required to accept international standards . . . that would result in a lower level of protection than the government has determined to be appropriate."¹⁰⁴ Instead, the SPS Agreement promotes harmonization by having members choose one of three alternatives.

First, a member may choose to *conform* its sanitary measures to existing international standards, guidelines, or recommendations.¹⁰⁵ For example, the SPS Agreement identifies the standards or guidelines established by the Codex Alimentarius Commission as the relevant international reference body for food safety.¹⁰⁶ The Codex Commission establishes "acceptable daily intakes" ("ADIs") for certain kinds of chemicals in food, and also recommends "maximum residue limits" ("MRLs") for certain kinds of residues in foods.¹⁰⁷ Where Codex has set an ADI and

102. See *supra* Part I.B.1. The SPS Agreement formulates the appropriate balance: "Desiring to further the use of harmonized sanitary and phytosanitary measures between Members, . . . without requiring Members to change their appropriate level of protection of human, animal or plant life or health" SPS Agreement, *supra* note 3, preamble.

103. See SPS Agreement, *supra* note 3, arts. 3.1-3.3.

104. Statement of Administrative Action, *supra* note 55, § 7.

105. See SPS Agreement, *supra* note 3, art. 3.2.

106. SPS Agreement, *supra* note 3, annex A.3(a); Hormones Panel Report (CAN), *supra* note 13, ¶ 8.62.

107. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.62. See generally FAO/WHO Consultation on Risk Management, *supra* note 37; FAO/WHO Expert Consultation on Risk Analysis, *supra* note 24.

Existing Codex standards reflect what Codex considers to be an acceptable level of protection. See FAO/WHO Consultation on Risk Management, *supra* note 37, at 4. The Joint Consultation stated:

[a]lthough industry and national regulators strive for production and processing systems which ensure that all food be "safe and wholesome," complete freedom from risks is an unattainable goal. Safety and wholesomeness are related to a level of risk that society regards as *reasonable in the context*, and in comparison with other risks in everyday life.

A Codex standard is the minimum standard for a food elaborated by [the Codex Alimentarius Commission] "so as to ensure a sound, wholesome product" The word "minimum" does not have any pejorative connotations and simply

recommended an MRL, a member could adopt the ADI and choose to incorporate the MRL into its own domestic law as a legally permissible residue level. The SPS Agreement encourages such actions. It provides that a conforming sanitary measure would be "deemed to be necessary" for purposes of article 2.2 and would be "presumed to be consistent with" other relevant provisions of the SPS Agreement and of the General Agreement on Tariffs and Trade 1994 (GATT 1994).¹⁰⁸

Second, a member may choose to "base [its] sanitary . . . measures on international standards," instead of simply conforming to those standards.¹⁰⁹ For example, a member might adopt the ADI but adjust the MRL to reflect dietary patterns peculiar to its population, or other considerations involved in exposure assessment. Alternatively, a member might adopt the level of protection implicit in the ADI but establish a measure

means the level of quality and soundness of a product judged by consensus to be appropriate for trade internationally and nationally.

Id. (emphasis added) (footnote omitted). See also *id.* at 8 (enumerating examples of "risk assessment policies" employed by scientific advisory groups relied upon by Codex).

As the Hormones Panel noted, an ADI is an estimate of the amount of a veterinary drug "that can be ingested daily over a lifetime without appreciable health risk." Hormones Panel Report (CAN), *supra* note 13, ¶¶ 2.17, 8.64 (quoting from Residues of Veterinary Drugs in Foods, in JOINT FAO/WHO CODEX ALIMENTARIUS COMMISSION, 3 CODEX ALIMENTARIUS 65 (1995)). See *Codex Maximum Residue Limits for Pesticides* (visited May 7, 1998) <<http://www.fao.org/waicent/faostat/pest%2Dresidue/pest%2De.htm>> (defining the "ADI" for a chemical as "the daily intake which, during an entire lifetime, appears to be without appreciable risk to the health of the consumer on the basis of all the known facts at the time of the evaluation of the chemical by the Joint FAO/WHO Meeting on Pesticide Residues"). For example, in the case of zeranone, a synthetic hormone at issue in the Hormones Dispute, which Codex considered a weak oestrogen that mimics oestradiol-17β, the Joint FAO/WHO Expert Committee on Food Additives set an ADI by "[a]dopting what it considered to be a conservative approach by . . . using a safety factor of 100." Hormones Panel Report (CAN), *supra* note 13, ¶ 8.67. The use of safety factors is explicitly recognized as an example of risk assessment policy, and setting an ADI for a tumorigenic agent reflects a science policy decision by the Joint Expert Committee and by Codex. See FAO/WHO Consultation on Risk Management, *supra* note 37, at 5, 8, 15-18. A national regulation implementing an MRL would be "primarily a regulatory tool to ensure that intake does not exceed the ADI and that good practice is observed." Hormones Panel Report (CAN), *supra* note 13, ¶ 8.64. Thus, an ADI may reflect a value judgment about how much risk is "appreciable," especially with regard to carcinogens, and an MRL is clearly a recommendation for risk management, not risk assessment. Cf. *id.* ¶ 8.77 ("[T]he fact that an ADI or MRL can reflect a level of protection (without *stricto sensu* itself being a level of protection), does not exclude . . . that an ADI or MRL can also be a sanitary measure in the sense of the SPS Agreement.").

108. See SPS Agreement, *supra* note 3, art. 3.2. Cf. John S. Eldred & Shirley A. Coffield, *What Every Food Manufacturer Needs to Know: Realizing the Impact of Globalization on National Food Regulation*, 52 FOOD & DRUG L.J. 31, 33 (1997) ("The WTO dispute resolution structure gives 'teeth' to Codex commodity standards and other international standards by allowing governments to point to Codex standards as persuasive authority when trade disputes arise at the WTO."). General Agreement on Tariffs and Trade 1994, 108 Stat. 4809 (1994), 33 I.L.M. 1154 (entered into force Jan. 1, 1995).

109. SPS Agreement, *supra* note 3, art. 3.1 (emphasis added). Note that the meaning of "based on" in SPS article 3.1 is different from, but consistent with, the meanings of "based on" in articles 2.2 and 5.1. See *id.* arts. 2.2, 5.1. While the former relates a member's sanitary measures to Codex standards, the latter relate a member's sanitary measures to scientific principles and to risk assessment, respectively.

different than a permissible residue level as a means of achieving that level of protection, in effect adopting what amounts to an "equivalent measure" in the sense of SPS article 4.1. In either of these examples, the central meaning of "based on" seems to be that the sanitary measure established by the member is related to the Codex ADI by scientific principles and offers the same level of protection as the Codex standards.¹¹⁰

Third, a member may choose to establish sanitary measures that provide "a higher level of sanitary . . . protection" than would measures that are "based on" the Codex standards.¹¹¹ In such a circumstance, however, one of two conditions must be satisfied: either there must be a "scientific justification" for the sanitary measures, or the sanitary measures must be "a consequence of" a higher level of protection selected in accordance with the relevant provisions of SPS article 5.¹¹² Regardless of which of these conditions is met, measures that achieve a different level of protection than that achieved by international standards must be consistent with all other relevant provisions of the SPS Agreement.¹¹³

These two alternative and independently sufficient conditions under the third option present a difficult task of interpretation.¹¹⁴ Given the definition of "scientific justification" that is provided,¹¹⁵ there seems to be little distinction between the two conditions. The first focuses on a member's judgment about the validity of the scientific reasoning behind an international standard or on the adequacy of an internationally recom-

110. It is not clear why the Hormones Panel found that SPS article 3.2 "equates measures based on international standards with measures which conform to such standards," Hormones Panel Report (CAN), *supra* note 13, ¶ 8.75. No reason is given for this assertion, which denies the possibility that a measure can be based on an international standard and adopt the same level of protection, without conforming to a specific recommended measure as a means of achieving that level of protection. See Appellate Body Report, *supra* note 16, ¶ 171 ("Under Article 3.1 . . . an SPS measure that is based on the existing relevant international standard . . . may adopt some, not necessarily all, of the elements of the international standard.").

111. SPS Agreement, *supra* note 3, art. 3.3; Appellate Body Report, *supra* note 16, ¶ 172.

112. SPS Agreement, *supra* note 3, art. 3.3.

113. *Id.* As the U.S. has stated: "By contrast [with the presumption in favor of a conforming measure], the fact that a sanitary or phytosanitary measure differs from a relevant international standard, guideline, or recommendation does not, in itself, create any adverse presumption concerning that measure." Statement of Administrative Action, *supra* note 55, § 7. This is surely the correct view. The creation of a presumption for conforming standards under SPS article 3.2 should not imply any presumption against measures that are merely "based on" or are "higher than" the international standard.

114. See SPS Agreement, *supra* note 3, art. 3.3; Appellate Body Report, *supra* note 16, ¶ 175.

115. Note 2 to SPS article 3.3 reads:

For the purposes of paragraph 3 of Article 3, there is a scientific justification if, on the basis of an examination and evaluation of available scientific information in conformity with the relevant provisions of this Agreement, a Member determines that the relevant international standards, guidelines or recommendations are not sufficient to achieve its appropriate level of sanitary or phytosanitary protection.

mended means to actually achieve a given level of protection.¹¹⁶ The second condition focuses on the selection of a different level of protection by a member.¹¹⁷ In either case, a justification must consist of an assessment of the risk and a determination that the chosen measure does in fact bring about a risk reduction when compared with the international standards. In the end, the two conditions seem to add no substantive requirements that are not already found elsewhere in the Agreement. What then is the point of stating them as two separate and alternative conditions? Perhaps one reason is to emphasize a difference in the scope of review due under the two conditions. Stating the second condition separately emphasizes that a member's selection of a level of protection is largely unreviewable, whereas different standards apply when a risk assessment or a determination about the means of achieving a selected level of protection is reviewed. Although the SPS requirements for each of these functions are prescribed elsewhere in the Agreement,¹¹⁸ the separate statement of these two conditions in article 3.3 arguably emphasizes the distinction.

The separate statement of these two conditions in article 3.3 takes on added significance when we consider how this "soft" harmonization strategy of the SPS Agreement is likely to affect the harmonization of science policies by members. I argued above that the adoption of particular science policies to use in risk assessment should be viewed as an aspect of selecting an appropriate level of protection.¹¹⁹ This means that a more protective measure that a member justifies by its adopted science policies should be evaluated under the second condition of article 3.3. By contrast, a more protective measure that is justified by appeal to purely scientific reasoning (whether about toxicity, exposure, or means-end causality) should be evaluated on its merits as a "scientific justification" under the first condition. Therefore, the two conditions partition the possible types

Id. art. 3.3 n.2.

116. The first condition is merely that there be a "scientific justification," see SPS Agreement, *supra* note 3, art. 3.3, which is elucidated by Note 2, quoted *supra* note 115. At least one commentator has suggested that this footnote might blur the distinction between science and policy. Wirth, *supra* note 1, at 827. As I argue below, however, in the text accompanying *infra* note 120, this footnote and SPS article 3.3 can also be read as recognizing the distinction between science and policy.

According to the Hormones Panel, both Canada and the European Communities interpreted the first condition to be a possibility in situations when the international standard is "outdated" or "inadequate, faulty or obsolete from a scientific point of view," such as where the standard turns out not to provide the level of protection it had been thought to provide. See Hormones Panel Report (CAN), *supra* note 13, ¶ 8.84. However, a new sanitary measure might be "based on" the intended level of protection, yet not "conform to" the inadequate international recommendation or standard. This would help give the SPS Agreement the temporal flexibility essential in those regulatory areas characterized by pervasive scientific uncertainty but increasing scientific knowledge.

117. The second condition is if the more protective measure is "a consequence of the level of sanitary or phytosanitary protection a Member determines to be appropriate in accordance with the relevant provisions of paragraphs 1 through 8 of Article 5," which deal with conducting risk assessment and determining the appropriate level of sanitary or phytosanitary protection. See SPS Agreement, *supra* note 3, arts. 3.3, 5.1-5.8.

118. See *supra* Parts I.B.1, I.B.2.

119. See *supra* Part I.B.1.

of justification into "scientific" and "policy" categories.¹²⁰ If "scientific justifications" turn out to be more reviewable than policy decisions, then harmonization of the former can be promoted more vigorously than harmonization of the latter — a result that is consistent with the tenor of the SPS Agreement.

II. WTO Review of Sanitary Measures Established by Members

Under the SPS Agreement, every member retains the sovereign right to adopt those sanitary measures necessary to achieve the level of protection that the member selects as appropriate for its society. On the other hand, the Agreement seeks to ensure that sanitary measures that restrict trade are not discriminatory and are not in fact "disguised restrictions on international trade."¹²¹ The substantive and procedural provisions of the SPS Agreement are designed to accommodate both of these major interests.¹²²

In general, it is easy to appreciate some of what is required in order for the WTO to balance these interests successfully over the long run. Substantively, the WTO needs to understand the issues appropriate for factfinding by a panel. Factfinding panels must adequately understand the nature of risk assessment, the pervasiveness of scientific uncertainty, and the role of science policy. Panels must also respect the differences between risk assessment and risk management, as well as the interactions between them.

In terms of process, panel factfinding should be fair and equitable in its procedures, transparent in its findings and rationales, reasonably predictable in its methods and outcomes, and reasonably efficient.¹²³ More-

120. As we have seen, "risk assessment" contains both science and policy, see *supra* Part I.A, as does "risk management," see *supra* Part I.B. SPS article 3.3 is one place where the Agreement comes close to recognizing a distinction between science and policy. See SPS Agreement, *supra* note 3, art. 3.3.

121. See, e.g., *id.* arts. 2.3, 5.6.

122. Cf. Michele D. Carter, Note, *Selling Science Under the SPS Agreement: Accommodating Consumer Preference in the Growth Hormones Controversy*, 6 MINN. J. GLOBAL TRADE 625, 649 (1997) (stating that the "overall goal" of the SPS Agreement is "to lessen the burdens on international trade created by those health regulations not based on scientific evidence").

123. Cf. DSU, *supra* note 7, arts. 3.2, 3.3, appendix 3.10. The U.S. Congress has provided by statute that:

[t]he Trade Representative shall seek the adoption by the Ministerial Conference and General Council of procedures that will ensure broader application of the principle of transparency and clarification of the costs and benefits of trade policy actions, through the observance of open and equitable procedures in trade matters by the Ministerial Conference and the General Council, and by the dispute settlement panels and the Appellate Body under the Dispute Settlement Understanding.

Uruguay Round Agreements Act, 19 U.S.C. § 3536 (1994).

Cf. Gresham & Bloomfield, *supra* note 61, at 1160-62 ("[T]he DSU greatly improves public access to information in the dispute settlement process."); C. O'Neal Taylor, *The Limits of Economic Power: Section 301 and the World Trade Organization Dispute Settlement System*, 30 VAND. J. TRANSNAT'L L. 209, 315 (1997) (stating that with regard to alleged violations of international trade law, the WTO dispute settlement system is "both more legitimate and equitable than the self-help of unilateralism"); Kim Rubenstein &

over, the acceptability of the panel process rests upon the appropriateness of the standard governing WTO panels in reviewing the factual and policy determinations made by members, of the "burdens of proof" placed on parties, and of the appellate review over the panel findings.¹²⁴

It is easy enough to state such broad, abstract goals. However, when trying to work out the details, one encounters deeper conceptual problems. Fine-tuning these aspects of WTO factfinding provides the best hope for keeping the desire for trade efficiencies from threatening sovereignty, and thus the best hope for achieving a stable long-term balance. A dominant theme running through this Article is that a detailed understanding of scientific uncertainty and science policy, and of their proper treatment within the WTO factfinding process, provides an important means of achieving the desired long-term balance. Central to this task is appreciating why we must keep the WTO from becoming the "World Trans-science Organization," and determining how this can be accomplished.

A. The Proper Substantive Issues Before a Panel

This part analyzes the content of the findings of fact to be made by WTO panels. The focus of attention is on those findings that pose the greatest

Jenny Schultz, *Bringing Law and Order to International Trade: Administrative Law Principles and the GATT/WTO*, 11 ST. JOHN'S J. LEGAL COMMENT. 271, 305 (1996) (stating that the WTO dispute settlement process is a "clear movement toward administrative law principles of due process/natural justice, broader standing, more transparency, and an appeal mechanism which emphasizes some form of judicial review"); Michael K. Young, *Dispute Resolution in the Uruguay Round: Lawyers Triumph Over Diplomats*, 29 INT'L LAW. 389, 409 (1995) (stating that "[c]onsistency, coherence, and persuasiveness" are qualities of substantive action necessary to confer legitimacy, as is procedural fairness).

124. Professor Wirth observed the following problems facing the WTO dispute settlement process:

A number of singular issues arise from the likelihood, given the structure of the new trade disciplines in the Uruguay Round and the NAFTA texts on standards, that the quasi-adjudicatory dispute settlement panels will be obliged to review the scientific foundation for national regulatory measures. The four texts expressly address none of the following three central questions:

- To what extent, if at all, must panels defer to expert scientific judgment underlying a national standard, especially if that judgment reflects minority or controversial views within the scientific community?
- To what extent, if at all, must panels defer to expert scientific judgment underlying a national standard when that judgment is exercised under conditions of scientific uncertainty?
- How should dispute settlement panels treat and structure requests for expert scientific advice in addressing issues raised by the previous two questions?

Wirth, *supra* note 1, at 853. Wirth concluded that the solution must lie in the direction of deference:

[t]he structure of these texts, the best thinking on the role of science in the national regulatory process, considerations of scientific uncertainty, and the negotiating histories of these agreements all suggest that dispute settlement panels should be highly deferential to scientific determinations of national authorities that underlie regulatory measures to protect the environment and public health.

Id. at 858-59. However, an adequate solution is far more complicated than the deference that Wirth proposed.

risk that a panel will overlook scientific uncertainty and science policy and will confuse risk assessment with risk management.

1. Reviewing Risk Assessment Determinations

A primary issue before a WTO panel is whether a member's measures are "based on" a risk assessment appropriate to the circumstances, are "based on" scientific principles, and are maintained with "sufficient scientific evidence."¹²⁵ In short, the primary issue of fact before a panel should be whether there is any reasonable scientific basis for a member's sanitary measures. The United States has stated: "[i]t is clear that the requirement in the [SPS] Agreement that measures be based on scientific principles and not be maintained 'without sufficient scientific evidence' would *not* authorize a dispute settlement panel to substitute its scientific judgment for that of the government maintaining the sanitary or phytosanitary measure."¹²⁶ For example, in reviewing a member's determinations in the context of risk assessment, a panel should not decide whether it is *true* that the contested food additive is a human carcinogen or whether it is *true* that there is no threshold for the carcinogenic action of the contested food additive. Instead, a panel's proper task is to make findings about whether there is *any reasonable scientific basis* for a member's determinations and sanitary

125. See *supra* Part I.A.

126. Statement of Administrative Action, *supra* note 55, § 6.b. As the Hormones Panel correctly concluded, the panel was not authorized to conduct its own risk assessment. Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.104, 8.118. But see John H. Barton, *Biotechnology, the Environment, and International Agricultural Trade*, 9 GEO. INT'L ENVTL. L. REV. 95, 103 (1996) ("[T]he chances are good" that WTO panels will be less deferential than the U.S. executive branch hopes, and will determine that "a phytosanitary standard must be based on scientific principles.").

The Appellate Body took a deferential approach in requiring *prima facie* evidence before shifting the burden of proof to a defending member in a dispute proceeding concerning the Agreement on Textiles and Clothing, see Report of the Appellate Body, United States — Measure Affecting Imports of Woven Wool Shirts and Blouses from India, Apr. 25, 1997, adopted May 23, 1997, 12-17, WT/DS33/AB/R, available in Westlaw, WTO-DEC file, 1997 WL 222239 [hereinafter Measure Affecting Imports of Woven Wool Shirts]; cf. Edward Krauland et al., *International Legal Developments in Review: 1996 Business Regulation, International Trade*, 31 INT'L LAW. 433, 441-42 (1997) (stating that while neither the Agreement on Textiles and Clothing nor the DSB's rules prescribe a specific standard of review for panels, some panels have "adopted a standard of review that is analogous to that applied by U.S. courts reviewing U.S. administrative agency determinations").

For a pre-WTO GATT panel discussion on standard of review, see Korea — Anti-Dumping Duties on Imports of Polyacetal Resins from the United States, April 2, 1993, adopted April 27, 1993, ¶ 227, ADP/92 (stating that a review to decide whether a determination by the Korean Trade Commission "was based on positive evidence did not mean that the Panel should substitute its own judgement . . . as to the relative weight to be accorded to the facts . . ."). As one commentator has noted, however, GATT 1947 contains no reference to science and no express requirement for deference to determinations by scientific experts. Wirth, *supra* note 1, at 845. Additionally, at least one pre-WTO panel is noteworthy for "its intrusive review of the exercise of expert scientific judgment by national regulatory authorities." *Id.* (referring to a panel applying GATT principles to Canada's landing requirement for Pacific Coast salmon and herring under the U.S.-Canada Free Trade Agreement).

measures.¹²⁷

On the other hand, a WTO panel cannot merely defer in its factfinding to any member that cries "science." Not only would this render the SPS Agreement ineffective as a trade agreement, but it could also perversely encourage global fragmentation in science by encouraging trade protectionist interests to co-opt the academy.¹²⁸ Both trade and science will benefit if the WTO conceptualizes its role as an independent factfinding body. If carefully balanced, this factfinding role can and should respect the important science-policy divide within risk assessment.

A WTO panel should make findings on whether, with respect to a particular issue, there is a scientific consensus or scientific uncertainty. Scientific uncertainty is evidenced by a good-faith difference of opinion about plausible accounts within the scientific community. Panels will usually discover that there is a consensus about the identity of the relevant sub-issues within risk assessment, about which methodologies and theories are beyond reasonable dispute, and about which models, hypotheses, or assumptions are so unreasonable as to be scientifically implausible. On the other hand, panels will also discover that, with regard to some distinct sub-issues, there is uncertainty about which models to employ or which input assumptions to make.¹²⁹ When a panel encounters such uncertainty, it should make findings concerning which alternative accounts scientists find plausible and which they do not. In so doing, a panel would map the contours of the current scientific terrain, determining where scientists agree on what is "proved" and what is "fanciful," as well as where scientists disagree over plausible alternative accounts. Whenever a panel finds a good-faith difference of opinion among scientists and finds any reputable scientific support for a member's science-policy choices within risk assessment, then the panel should find that the member's sanitary measures are "based on" scientific principles and have "sufficient scientific evidence."¹³⁰

127. See *supra* Part I.A; cf. Wirth, *supra* note 1, at 854-58 (Although "the Uruguay Round and the NAFTA texts on standards contain no express instruction that dispute settlement panels must accord scientific determination by national regulatory authorities some measure of deference," such a conclusion "is virtually inescapable.").

128. Cf. Atik, *supra* note 61, at 748-51, 757-58 ("We can anticipate many cases where scientific consensus is split along national lines. . . . Science . . . promises little hope as a source for neutral principles to resolve economic disputes among nations.").

129. Examples of such sub-issues are enumerated *supra* Part I.A.2.

130. As the Appellate Body stated:

Article 5.1 [of the SPS Agreement] does not require that the risk assessment must necessarily embody only the view of a majority of the relevant scientific community. In some cases, the very existence of divergent views presented by qualified scientists who have investigated the particular issue at hand may indicate a state of scientific uncertainty.

Appellate Body Report, *supra* note 16, ¶ 194. The Appellate Body did not establish any quantitative or qualitative standards for evaluating the extent of scientific support required under the SPS Agreement, holding instead that any determination should be made "on a case-to-case basis." *Id.*

The United States stated the standard in terms familiar within U.S. law: a panel's task is to find whether the member's measures are "based on scientific principles (rather

There are numerous policy arguments why WTO panels should make findings only about the reasonableness or plausibility of any member's risk assessment determinations. That is, there are good policy reasons why, in reviewing risk assessment determinations, a WTO panel should leave undisturbed the science-policy choices of a member, so long as that member's inferences from the available data are scientifically plausible.

First, the goal of true economic efficiency can be achieved only if the scientific understanding of risks, costs, and benefits is accurate. The process of scientific inquiry, however, is usually incomplete, is almost always open to revision, and often results in differences among scientific opinions. Therefore, when scientists themselves are uncertain which of two alternative accounts is the correct or accurate one, and members choose differently from among these plausible accounts, then the WTO should wait for the scientific community to resolve that uncertainty, rather than impose the views of a single group of scientists on all WTO members.¹³¹

WTO panels, therefore, should determine whether there is a scientific consensus on a factual issue and whether there is a good-faith difference of opinion among scientists. Whenever a panel finds scientific uncertainty and finds a difference of opinion among reputable scientists, then the panel should not impose its own opinion of the matter on members. The existence of scientific uncertainty itself should be compelling evidence to a WTO panel that each of the alternative accounts under debate is scientifically plausible, and that a sanitary measure justified by any such account is "based on" scientific principles and has "sufficient scientific evidence."¹³² WTO panels should respect the open-endedness of scientific inquiry. If they do not, they will be abandoning "good science" as the foundational principle of the SPS Agreement, and placing a higher value on

than, for instance, requiring measures to be based on the 'best' science)" and whether there is "sufficient scientific evidence (rather than, for instance, requiring an examination of the 'weight of evidence')." Statement of Administrative Action, *supra* note 55, § 6.b.

Some commentators have urged an analogy to the law governing judicial review of the scientific determinations of administrative agencies such as the EPA or FDA. This aspect of U.S. administrative law is a carefully constructed balance between a court's duty to ensure administrative transparency and the rule of law, on the one hand, and the expertise and policy-setting mandate of the administrative agencies, on the other. See, e.g., *Ethyl Corp. v. EPA*, 541 F.2d 1, 36-37 (D.C. Cir. 1976), *cert. denied*, 426 U.S. 941 (1976) (ruling that the pertinent arbitrary and capricious standard is merely an inquiry into whether the agency decision at issue was rational). However, several distinguishing characteristics between WTO panels and domestic reviewing courts are discussed in this part of the Article and *infra* Part II.B, and the analogy should not be pressed too far. One distinction is that WTO panels are factfinding entities of an international trade institution, the findings of which are subject to only limited appellate review. The constitutional matrix of the U.S. government cannot be simply transferred to the international setting.

131. But see Carter, *supra* note 122, at 648-49 (The fact that "there is no accepted level of 'certainty' with regard to scientific data . . . directly conflicts with the stated goal of the SPS Agreement. Moreover, in order to regulate confidently, Members must be given a baseline level of certainty which can be used to justify health-related regulations.").

132. See SPS Agreement, *supra* note 3, art. 2.2.

WTO-imposed uniformity than on regulation that is genuinely efficient and is based on accurate science.

Second, a panel's findings of fact should not undermine the integrity and autonomy of the members' domestic regulation of human health and safety.¹³³ The risk assessment scientists within the regulatory institutions of a member state should not have to worry about their scientific determinations being second-guessed, years after the fact, by a WTO panel deciding an international trade dispute. Restricting WTO factfinding to "reasonableness" should provide enough predictability to safeguard domestic regulatory factfinding from overly intrusive WTO factfinding. Moreover, findings of fact on scientific issues often have implications beyond a single dispute, creating precedent for science-policy determinations in legally dissimilar areas. For example, a major impetus behind the development of risk assessment methodology in the United States was a desire to harmonize different agencies' approaches to regulating carcinogenicity.¹³⁴ A WTO panel decision that a particular compound is not carcinogenic might indirectly affect a purely domestic controversy about the appropriate level of cleanup for a toxic waste site. If WTO panels are empowered to make choices among scientifically plausible alternatives and to impose those choices on members, then this may affect risk assessment in domestic regulatory areas unrelated to international trade.

Third, as discussed above, most risk assessments incorporate science policies into decisions about which factual assumptions to make in the face of data gaps, or about which mathematical models to use in dose-response assessment or exposure assessment.¹³⁵ Science-policy choices are pervasive and numerous. It can be extremely difficult to identify every assumption and default rule underlying a given risk assessment, and the adversarial context of a WTO panel is not the ideal forum in which to inventory all science policies. Moreover, as argued above, WTO panels should have only limited authority to find that a member's decision to adopt a particular science policy is inconsistent with the SPS Agreement.¹³⁶ In view of these major constraints on legitimate WTO factfinding, it would be prudent to restrict factfinding to a zone of reasonableness when risk assessment determinations involve scientific uncertainty.

Fourth, a WTO panel does not have the institutional capacity to decide among scientifically plausible alternatives. Domestic regulatory agencies such as EPA or the European Commission have the depth and continuity of expert staff needed to weigh the conflicting evidence about scientific uncertainty. These agencies can also make a weight-of-evidence determination at any one point in time, and stand ready to re-evaluate that evidence when new studies become available. A WTO panel, on the other

133. Cf. Wirth, *supra* note 1, at 854-58 (arguing that concern about disrupting national regulatory programs is one factor in favor of showing deference to national scientific determinations).

134. See, e.g., NRC (1983), *supra* note 21, at 1-8.

135. See *supra* Part I.A.

136. See *supra* Part I.B.1.

hand, is formed to decide the merits of a particular case at a particular point in time. It has neither the resources nor the mandate to monitor and re-evaluate an evolving scientific debate. A WTO panel should leave any member's scientific judgments undisturbed whenever the panel finds that the current state of science provides a reasonable basis for those judgments.

Fifth, appellate review within the WTO is limited. If WTO panels are empowered to evaluate the relative merits of plausible scientific alternatives and to choose which alternative to impose upon a party to a WTO dispute, then this "finding of fact" by the panel cannot be reviewed effectively by the Appellate Body.¹³⁷ Therefore, over time, different WTO panels may make different and inconsistent scientific determinations. This could have a confounding effect both on the scientific debate itself and on the domestic regulatory institutions of WTO members. The only viable solution to this problem is to hold that WTO panels do not have the authority to decide which plausible scientific alternative is most likely to be true. Panels should have authority only to decide whether there is any reputable support within the scientific community for a member's determination of scientific fact.¹³⁸

Sixth, whenever the subject matter of a trade dispute is the protection of human health, there is an additional policy reason for deferring to those plausible determinations by members that tend to protect human beings from additional risk.¹³⁹ When there is a reasonable scientific basis for a member's determination that a sanitary measure will provide better protection for human health, then there is no compelling and competing WTO policy that justifies disregarding scientific uncertainty or resolving a scientific dispute by fiat.

Seventh, although the WTO has a legitimate goal of fostering the transparency of members' factfinding when a member's sanitary measures

137. See DSU, *supra* note 7, art. 17.6 ("An appeal shall be limited to issues of law covered in the panel report and legal interpretations developed by the panel.").

138. See Wirth, *supra* note 1, at 855 (arguing that "any domestic scientific determination that has withstood scientific peer review should be categorically presumed to satisfy the science-based disciplines in either the Uruguay Round or the NAFTA"). What can pass for "peer review," however, is so diverse that it should not be considered, as Wirth claims, "the scientific analogue of a procedural . . . test." *Id.* Wirth himself acknowledges that there are possible weaknesses in "peer review." *Id.* at 842-43.

139. See Appellate Body Report, *supra* note 16, ¶ 194 (stating that the fact of members acting in good faith on the basis of scientific opinion that diverges from the "mainstream" scientific opinion "does not necessarily signal the absence of a reasonable relationship between the SPS measure and the risk assessment, especially where the risk involved is life-threatening in character and is perceived to constitute a clear and imminent threat to public health and safety"). *Cf.* Wirth, *supra* note 1, at 851 (noting that in the panel reports in cases prior to the Uruguay Round and NAFTA, several cases "may suggest a predisposition to greater deference in the context of regulations designed to protect public health . . . than in situations involving preservation of the environment or natural resources . . ."). A failure by the WTO to give added deference might itself constitute a disregard for the precautionary principle of international environmental law. See *supra* note 77.

restrict the flow of trade,¹⁴⁰ this goal can be achieved by requiring members to disclose their reasons for establishing those sanitary measures.¹⁴¹ Once the scientific bases for those underlying factual determinations have been made clear to the international community, whether in proceedings before the domestic regulatory institutions of a member or in scientific evidence produced before the WTO panel, this policy objective has been achieved. Therefore, the objective of fostering transparency does not itself justify WTO factfinding about the truth of a scientific issue, as contrasted with factfinding merely about reasonableness. On the contrary, the interests of transparency are not served by having WTO factfinding panels behaving like quasi-scientific expert panels, mimicking scientific investigations, and making findings of fact about the truth of scientific propositions. WTO panels can better promote transparency by acting merely as careful consumers of scientific information. WTO panels should remain generalist and non-expert factfinders, who hear the scientific explanations that the parties and expert advisors present to them, and make findings only on the issue of the "reasonableness" of those explanations.

Finally, the WTO may have a legitimate goal of promoting domestic regulatory procedures that allow potentially affected private parties or potentially affected members to produce scientific evidence relevant to particular sanitary measures and to ensure consideration of that evidence by the member adopting the sanitary measures.¹⁴² Once members have had an opportunity to produce scientific evidence, however, whether in domestic regulatory proceedings or before a WTO panel, this policy objective has been achieved. Therefore, this objective does not require WTO factfinding about the relative weight of evidence among plausible scientific accounts.

If this analysis is correct, and WTO panels should engage in factfinding only about the reasonableness of a risk assessment or about the plausibility of scientific determinations made by members in the context of risk assessment, a consensus on "scientific plausibility" itself may evolve.¹⁴³ I propose that WTO panels should make findings about the extent of scientific uncertainty in risk assessments, and about which alternative accounts have sufficient backing by reputable scientists. As panels explain their rationales for such findings, they may develop criteria for scientific plausi-

140. See, e.g., SPS Agreement, *supra* note 3, art. 7; DSU, *supra* note 7, app. 3.10; cf. Phillip R. Trimble, *Globalization, International Institutions, and the Erosion of National Sovereignty and Democracy*, 95 MICH. L. REV. 1944, 1968 (1997) (reviewing THOMAS M. FRANCK, *FAIRNESS IN INTERNATIONAL LAW AND INSTITUTIONS* (1995)) ("A corollary to more open participation at the norm-creation stage, and a key to accountability, is that the decisionmaking process must be more transparent.").

141. SPS Agreement, *supra* note 3, art. 5.8.

142. This may be a long-range result deriving indirectly from the WTO dispute settlement process, rather than an explicitly stated goal. This parallels American law, where the need for a record upon which to base judicial review has prompted agencies to create administrative records where none are explicitly required by statute. See *Citizens to Preserve Overton Park, Inc. v. Volpe*, 401 U.S. 402, 416-21, 91 S. Ct. 814, 823-26 (1971).

143. Cf. Wirth, *supra* note 1, at 855-56 (suggesting a list of questions that a panel might consider "in determining whether a challenged measure qualifies as minimally 'scientific' in situations where prior peer review is absent").

bility that are acceptable not only to members and interested private parties, but also to the scientific community itself. If such criteria evolve, then WTO proceedings can become more objective and more predictable, and the potential for WTO intrusion into domestic risk assessment will remain limited.

2. Reviewing Risk Management Decisions

As discussed above, a member's selection of a level of protection is a value-based act of sovereignty and is subject to very limited review.¹⁴⁴ In reviewing such a selection at the present time, a WTO panel need only find the mere fact of selection. Perhaps the standing Committee will develop guidelines for determining whether different internal levels of protection are "consistent," or whether they represent "arbitrary and unjustifiable distinctions."¹⁴⁵ Even when different levels are selected by a member, the difference might be "justified" by considering the balance of risks and benefits.¹⁴⁶

The WTO has a greater factfinding role with respect to the measures chosen as means to achieve a selected level of protection. First, a panel considers whether a given sanitary measure is "based on" a risk assessment appropriate to the circumstances, is "based on" scientific principles, and is supported by "sufficient scientific evidence."¹⁴⁷ These findings are simply those discussed in the previous part on risk assessment.¹⁴⁸ The issue to be decided is whether there is any reasonable scientific basis for a member's assessment of the risk.

In evaluating measures as means, however, science can play an evidentiary role in issues beyond risk assessment. A panel should also determine whether there is a reasonable cause-and-effect relationship between the sanitary measure chosen and the member's selected level of protection.¹⁴⁹ A trade-restrictive sanitary measure is justifiable only if it is reasonably effective in bringing about actual protection – in reducing or avoiding the risks identified by risk assessment. Also at issue is whether there is an alternative measure that is at least as effective at risk reduction as the measure chosen, but which has collateral effects that are less trade-restrictive.¹⁵⁰ Each of these questions is factual and involves determinations of causation. They are therefore proper objects of scientific inquiry. In fact, they are usually questions about differences in resulting exposure under

144. See *supra* Part I.B.1.

145. See *id.*

146. It is beyond the scope of this Article to propose how these concepts should be elucidated. However, it is worthwhile to note here that cost-benefit rationales for different levels of protection will usually rest on causal and economic models that employ both data and input assumptions. These rationales will also contain scientific uncertainty and may employ science policies. They may therefore require review by WTO panels similar to the review of risk assessments. That is, assessments of expected net benefits should pass scrutiny if but only if they are scientifically plausible.

147. See SPS Agreement, *supra* note 3, arts. 5.1, 2.2.

148. See *supra* Part II.A.1.

149. See *supra* Part I.B.2.

150. See *id.*

the various measures being compared, and can be answered systematically only by conducting exposure assessment and risk assessment. Such comparisons also involve questions of comparative efficiency, and the extent to which the measures impose different burdens on international trade.

Scientific determinations are also involved whenever a panel must decide whether an exporting member's means of protection is "equivalent" to an importing member's sanitary measures. Again, the issue presented is whether alternative measures would achieve the same level of protection. Making such a determination requires conducting risk assessment for each alternative measure and comparing the results.

Many of the same issues are presented when a complaining member alleges a violation of SPS article 3, which deals with harmonization.¹⁵¹ As discussed above,¹⁵² a member must select one of three options: (1) it may conform its sanitary measures to existing international standards, guidelines or recommendations; (2) it may base its sanitary measures on such international standards but not merely conform to them; or (3) it may adopt measures that provide a higher level of protection than would measures based on such international standards. Findings of conformity would presumably be the easiest for a panel to make. If option (2) is taken, however, a panel would be called upon to review the scientific justification for the alternative, non-conforming measure. If option (3) is at issue, a panel must compare risk assessments to determine whether the measure chosen is in fact more protective than a conforming measure would be. Thus, the contents of the findings required to resolve disputes under article 3 of the SPS Agreement are of the same type as those for other disputes involving risk assessment and risk management.

The important thing to note is that in reviewing the risk management decisions of a member, WTO panels are repeatedly called upon to review risk assessment and scientific determinations by members. Therefore, the argument developed above also applies here: the task of a WTO panel should be to find whether there is any reasonable basis for a member's risk assessment and scientific determinations, and whether the member's science-policy choices employed scientifically plausible alternatives.¹⁵³ This formulation of the proper substantive issues before a WTO panel recognizes the panel's objective factfinding role, but keeps a panel from second-guessing both scientists and the science policy officials of members. It also affords the deference required when science policies implement a sovereign member's selection of an appropriate level of protection.

151. See SPS Agreement, *supra* note 3, art. 3.

152. See *supra* Part I.C.

153. See *supra* Part II.A.1. Cf. *Industrial Union Dep't, AFL-CIO v. American Petroleum Inst.*, 448 U.S. 607, 656, 100 S. Ct. 2844, 2871 (1980) (plurality opinion) (stating that, under the federal Occupational Safety and Health Act, "so long as they are supported by a body of reputable scientific thought, the Agency is free to use conservative assumptions in interpreting the data with respect to carcinogens, risking error on the side of overprotection rather than underprotection").

B. The Evidence Available to a Panel

If WTO panels are to make findings about the scientific plausibility of the alternative accounts that are subject to science-policy choices by WTO members, what evidence should a panel consider in making its findings? The growth hormones dispute raised the issue of whether the Panel's deliberations should proceed on the basis of all evidence produced in its proceeding, or whether the Panel should confine itself to considering only the evidence that had been available to decision-makers prior to the proceeding.¹⁵⁴ The Hormones Panel decided to limit itself to the historical evidence, and to discount or ignore "new evidence" produced in the WTO proceeding by the European Communities.¹⁵⁵ A closely related question is whether to interpret the SPS requirement for "risk assessment" to require every member to undertake a risk assessment procedure prior to instituting a sanitary measure. If there is such a procedural requirement, then in any dispute about whether a measure is "based on an assessment . . . of the risks," the panel can regard the issue before it as an issue of historical fact. The panel would then be called upon to determine whether the defendant member had in fact conducted an adequate risk assessment prior to instituting or maintaining its challenged sanitary measure.

In possible conflict with this objective of enforcing procedural requirements on members, a panel has a substantive task. The mandate of a panel is to determine whether there is any reasonable scientific basis for a member's sanitary measures. The presumption should be that, to perform this task as well as possible, a panel should consider all relevant evidence reasonably available to the panel. This presumption should be overcome only when a competing and compelling institutional policy requires excluding "new evidence" from consideration.

WTO panels might be tempted to make their task more manageable by adopting a model of judicial review familiar from American administrative law in which the reviewing court confines its determination to the record that was before the agency at the time of the agency's decision.¹⁵⁶ However, WTO panels differ from domestic courts conducting judicial review of domestic administrative agencies. In the American system, the reviewing court is sometimes an appellate court, which has extremely limited

154. See Hormones Panel Report (CAN), *supra* note 13, ¶ 8.118. Such prior evidence might include the evidence available to the European Communities when they established the measures in dispute, or that available to the Codex Commission when it established the relevant ADIs and MRLs.

155. *Id.* This decision is discussed *infra* Part III.B.

156. See *Citizens to Preserve Overton Park, Inc. v. Volpe*, 401 U.S. 402, 419-20, 91 S. Ct. 814, 825 (1971); *Camp v. Pitts*, 411 U.S. 138, 142-43, 93 S. Ct. 1241, 1244 (1973) ("[T]he focal point for judicial review should be the administrative record already in existence, not some new record made initially in the reviewing court.").

Within American law, however, there are many institutional and constitutional policies bearing on this issue — policies such as separation of powers, autonomy of agencies, judicial integrity, and legislative intent (statutory interpretation). How similar policies would apply in the WTO context raises fundamental issues about institutional legitimacy and prudential policies that go far beyond the scope of this Article. The discussion here is limited, therefore, to considerations directly related to WTO factfinding.

factfinding authority. A WTO panel, by contrast, has primarily a factfinding mandate. In the American system, the issue before a reviewing court is often whether an agency acted or made findings in a manner that was "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law."¹⁵⁷ Other portions of the American statutory matrix explicitly direct courts to review how the agency acted in light of the record before the agency.¹⁵⁸ A WTO panel, by contrast, is charged with finding whether there is any reasonable scientific basis for a sanitary measure.¹⁵⁹ It should focus on the sanitary measure and the state of scientific knowledge at the time of the panel proceeding. While members' past domestic proceedings may be a valuable source of evidence, they should not provide the only evidence of scientific plausibility that a panel may consider.

A number of policy arguments can be given for why WTO panels should rely upon the best scientific information available. First, the WTO has an institutional interest in efficient regulation based on accurate science. Real efficiencies are likely to be created only if the cause-and-effect relationships surrounding products and trade are understood and respected. Domestic courts may be free to distance themselves from the ultimate effects of the regulatory decisions they are reviewing because the regulatory responsibility lies with the agency and the task of the court is merely to keep the agency from unlawful action.¹⁶⁰ But the WTO should not so distance itself. Realizing true economic efficiencies, redressing market imperfections, and achieving regulatory legitimacy depend upon accurately assessing scientific plausibility. Due to its interest in the efficient working of the international trade system, the WTO is more like an administrative agency than it is like a reviewing court.

Second, although the WTO has an interest in accurate science, the process of scientific inquiry often generates differences among scientific opinions. Thus, the WTO's institutional interest has two aspects. On the one hand, whenever scientists are uncertain, the WTO should wait for the scientific community to resolve that uncertainty. It should not impose one of the alternative plausible accounts on WTO members. On the other hand, the WTO should not blind itself to advances in scientific knowledge and the gradual resolution of scientific uncertainties. A WTO panel should consider whatever relevant evidence is available at the time of the dispute. In many cases, we can expect that the object of challenge will be

157. Administrative Procedure Act, 5 U.S.C. § 706(2)(A) (1994).

158. See, e.g., Administrative Procedure Act, 5 U.S.C. §§ 556, 557, 706(2)(E) (1994) (directing courts to review whether an agency's findings had substantial evidence in a record created at an administrative hearing).

159. See SPS Agreement, *supra* note 3, art. 2.2.

160. See, e.g., *Chevron U.S.A. Inc. v. Nat. Resources Def. Council, Inc.*, 467 U.S. 837, 864-66, 104 S. Ct. 2778, 2792-93 (1984) (holding that certain "policy arguments are more properly addressed to legislators or administrators, not to judges," that federal judges — who are not experts, are not part of either political branch of government, and have no constituency — "have a duty to respect legitimate policy choices" of the political branches, and that "[t]he responsibilities for assessing the wisdom of such policy choices . . . are not judicial ones").

old sanitary measures based on once accepted but now outmoded science. In such a case, a panel should be free to take into account recent evidence that shows an earlier view to be no longer plausible.

Third, the WTO has an interest in ensuring both that potentially affected members have an opportunity to produce scientific evidence relevant to particular sanitary measures, and that such evidence is properly considered.¹⁶¹ But the opportunity to produce scientific evidence can be assured only if it is afforded before a WTO panel. In the growth hormones case, the responding party wanted the panel to consider new evidence.¹⁶² It is also possible that the complaining party would seek to produce new evidence because it is dissatisfied with the defending party's domestic risk assessment proceedings. All parties to a dispute should have the right to present all relevant evidence to the panel and to have such evidence considered as fully as any other evidence.

Fourth, whenever there is uncertainty concerning the protection of human health, there is an additional policy reason for making a decision on the basis of the best scientific evidence available. When new evidence might present a reasonable scientific basis for a member's determination that a sanitary measure will provide better protection for human health, there is no compelling and competing WTO policy that justifies a panel's refusal to consider that evidence.

A fifth consideration is that WTO panels do not have the capacity to monitor scientific evidence, and it is costly to re-evaluate such evidence whenever new information becomes available. If a sanitary measure is challenged and a panel may consider only the scientific evidence available at the time the sanitary measure was adopted, a losing defendant member might simply conduct a new risk assessment and enact a new measure immediately after the DSB decision. Consequently, a new panel would be required to consider the new scientific evidence. It is much more efficient to review all of the scientific evidence available at the time the panel decides the merits of a particular case.¹⁶³ In this respect, a panel should act more like a court conducting a trial *de novo* than like a court conducting judicial review of an administrative agency's proceedings.

Finally, American courts have found that attempting to restrict judicial review to the record before the agency at the time of the agency's decision is a complicated task.¹⁶⁴ The problem of identifying what constitutes the agency's record is much easier in formal, trial-type administrative proceedings than it is in informal rulemakings.¹⁶⁵ Is the WTO to identify "the

161. See *supra* text accompanying note 142.

162. See *Hormones Panel Report (CAN)*, *supra* note 13, ¶ 8.118.

163. Cf. Dunoff, *supra* note 61, at 1111-13 (discussing the importance of impartial monitoring of changing scientific knowledge in the context of trade/environment conflicts).

164. See generally Steven Stark & Sarah Wald, *Setting No Records: The Failed Attempts to Limit the Record in Review of Administrative Action*, 36 ADMIN. L. REV. 333 (1984).

165. See *id.* at 338-39; KENNETH CULP DAVIS & RICHARD J. PIERCE, JR., 1 ADMINISTRATIVE LAW TREATISE § 7.4, at 313-14 (3d ed. 1994); ALFRED C. AMAN, JR. & WILLIAM T. MAYTON, ADMINISTRATIVE LAW § 13.10.3, at 500-04 (1993).

official record" for all of the different types of governmental proceedings undertaken by its members? Would such an approach lead to WTO rulings on what constitutes "the record for WTO purposes"? This approach seems both unwarranted and difficult. Moreover, even if a workable concept of such a record could be achieved, it is likely that WTO panels, like American courts, would not respect the boundaries set by their own doctrine.¹⁶⁶ For many of the policy reasons discussed above, panels will sometimes find it necessary to consider information not within the member's historical record. This should not be too surprising when one considers that a member's domestic record can be created at the member's discretion. A WTO panel should be free to perform its factfinding without being constrained by the record-making procedures of its members.

In conclusion, a WTO panel should consider all relevant scientific evidence available to it when it determines whether sanitary measures are supported by scientifically plausible accounts of risk.

C. The Standard and Burden of Proof

This part discusses the standard of proof that a panel should employ in its factfinding and other aspects of the burden of proof placed on parties.¹⁶⁷ It is essential to keep the following concepts distinct: (1) the standard of proof to be employed by the panel as it finds facts, (2) the burden of persuasion placed on parties, (3) the minimum requirements of rational inference and (4) the burden of producing evidence. These concepts apply to very different aspects of factfinding and respond to very different policies.

First, the standard of proof is the legal standard a panel should use to make its findings of fact.¹⁶⁸ The standard of proof establishes the quality of evidence and degree of certainty needed before a positive finding may be made. It is common to express the standard of proof appropriate in a non-criminal proceeding as the "preponderance" or the "greater weight" of the evidence.¹⁶⁹ A WTO panel should likewise assess all the evidence for and against any factual proposition and should adopt the proposition as a find-

166. See generally Stark & Wald, *supra* note 164 (cataloguing exceptions to the rule that judicial review should proceed on the basis of the administrative record alone).

167. Cf. Dunoff, *supra* note 61, at 1079-80 (noting that in challenges to environmental measures under NAFTA, issue of burden of proof is unclear and "is unlikely to be clarified until an actual dispute between the parties is adjudicated").

168. Cf. *Concrete Pipe and Prods. of Cal., Inc. v. Const. Laborers Pension Trust for S. Cal.*, 508 U.S. 602, 622, 113 S. Ct. 2264, 2279 (1993) (The three customary standards of proof — "preponderance," "clear and convincing," and "beyond a reasonable doubt" — indicate degrees of certainty required; before finding for the party who has the burden of proof, the factfinder "must evaluate the raw evidence, finding it to be sufficiently reliable and sufficiently probative to demonstrate the truth of the asserted proposition with the requisite degree of certainty.").

169. See generally FLEMING JAMES, JR. ET AL., *CIVIL PROCEDURE* § 7.14 (1992); 2 MCCORMICK ON EVIDENCE § 339 (John William Strong ed., 1992); Vern R. Walker, *Preponderance, Probability and Warranted Factfinding*, 62 BROOK. L. REV. 1075 (1996). On the use of the preponderance standard in international tribunals, see MOJTABA KAZAZI, *BURDEN OF PROOF AND RELATED ISSUES: A STUDY ON EVIDENCE BEFORE INTERNATIONAL TRIBUNALS* 347-50 (1996) (If "the *prima facie* evidence produced by the proponent of the burden of proof is challenged by the other party, and the tribunal decides the case on the basis of

ing if, but only if, the greater weight of evidence supports the proposition. A panel need not be certain that the proposition is true, nor does the proposition need to be scientifically proven. What should be required is a determination by a panel that the proposition at issue is more likely to be true than false.

The standard of proof is distinct from the content of the issues to be proved. The content of the principal issue before a panel is whether there is a reasonable scientific basis for the challenged sanitary measures.¹⁷⁰ The standard of proof to be used in making this finding is a preponderance of the evidence. The panel need not be *certain* that there is a reasonable scientific basis for the sanitary measures, or that the science-policy choices of the member are scientifically plausible. After considering all of the evidence on plausibility, the panel should find a member's choices to be plausible if the panel concludes that, more likely than not, there is a reasonable scientific basis for the choice.¹⁷¹

Second, there must be a default rule for making a finding when the weight of evidence is in "equipoise" – that is, when it appears to a panel that the evidence for a proposition seems equal in weight to the evidence for its negation.¹⁷² Factfinders need a "tie-breaking rule" to make findings when the evidence favors neither side. In such cases, the party with the "burden of persuasion" is the party who should lose on the issue unless it persuades the panel to its view by a preponderance of the evidence.¹⁷³ Parties to a treaty should enjoy a presumption that they are complying with the treaty unless a complaining party proves otherwise.¹⁷⁴ The complaining party should bear the burden of persuasion on the issue of whether there has been a treaty violation,¹⁷⁵ and findings should be made

evaluation of evidence produced by both parties, then the preponderance of evidence is the proper term.").

170. See *supra* Part II.A.

171. The confusion of content (what is to be proved) with standard (degree of proof required) encourages the mistake of considering a WTO panel as a reviewing court, with the task of determining only whether a member acted reasonably in the past by considering only the record before that member at the time the member acted. See *supra* Part II.B. The preponderance standard does not require certainty, but that does not mean that a panel is deferring to the factfinding of a member. Rather, the panel conducts its own factfinding into the scientific plausibility of a member's risk characterization.

172. On the need for a default rule and on the burden of persuasion, see generally JAMES ET AL., *supra* note 169, § 7.13; MCCORMICK ON EVIDENCE, *supra* note 169, § 336.

173. See JAMES ET AL., *supra* note 169, § 7.13, at 338 ("A concept of the risk of nonpersuasion is inseparable from any system in which issues of fact are to be decided through rational deliberation on the basis of incomplete knowledge.").

174. Cf. Vienna Convention on the Law of Treaties, opened for signature May 23, 1969, art. 26, 8 I.L.M. 679, 690 ("Every treaty in force is binding upon the parties to it and must be performed by them in good faith."); RESTATEMENT (THIRD) OF THE FOREIGN RELATIONS LAW OF THE UNITED STATES § 321 (1987) ("Every international agreement in force is binding upon the parties to it and must be performed by them in good faith.").

175. See Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 16. See John J. Barceló III, *Product Standards to Protect the Local Environment – The GATT and the Uruguay Round Sanitary and Phytosanitary Agreement*, 27 CORNELL INT'L L.J. 755, 764, 774-75 (1994) (arguing that the burden of proof should rest on the challenging WTO member, not the respondent); cf. KAZAZI, *supra* note 169, at 53-117, 221-35 (discussing

against the complaining party when the evidence of a violation is in equipoise.¹⁷⁶ Therefore, a complaining member should have the burden of persuading a panel that there is no reasonable scientific basis for the defending member's sanitary measures. In the context of scientific uncertainty and risk assessment, this would usually mean proving that any risk-oriented reasoning invoked by the defending member is not scientifically plausible and has no real support in the scientific community.¹⁷⁷

Third, the WTO Appellate Body should hold panels to a standard of rational inference.¹⁷⁸ A panel's finding should be reversed if the record before the panel does not contain the minimal evidence that any reasonable person would consider necessary to support such a finding.¹⁷⁹ For

the rule of *actori incumbit probatio*, the rule that the party who alleges a fact must prove the truth of its claim).

176. As the Appellate Body has stated with respect to the Agreement on Textiles and Clothing:

we find it difficult, indeed, to see how any system of judicial settlement could work if it incorporated the proposition that the mere assertion of a claim might amount to proof. It is, thus, hardly surprising that various international tribunals, including the International Court of Justice, have generally and consistently accepted and applied the rule that the party who asserts a fact . . . is responsible for providing proof thereof.

Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 14.

177. See *supra* Part IIA.

178. The DSU prescribes the factfinding function of panels:

a panel should make an objective assessment of the matter before it, including an objective assessment of the facts of the case and the applicability of and conformity with the relevant covered agreements, and make such other findings as will assist the DSB in making the recommendations or in giving the rulings provided for in the covered agreements.

DSU, *supra* note 7, art. 11. With respect to the Appellate Body, "[a]n appeal shall be limited to issues of law covered in the panel report and legal interpretations developed by the panel." *Id.* art. 17.6. Additionally, "[t]he Appellate Body may uphold, modify or reverse the legal findings and conclusions of the panel." *Id.* art. 17.13. However, the Appellate Body "lacks the power to remand a dispute back to the panel." Taylor, *supra* note 123, at 256.

The Appellate Body's authority to hold panels to a minimal standard of rational inference derives from the DSU requirements that a panel make an "objective assessment" of the case and that the Appellate Body should decide "issues of law" and "legal interpretations," including an interpretation of what constitutes an "objective assessment." See DSU, *supra* note 7, arts. 11, 17.6. If a panel were to reach factual conclusions without a reasonable basis in the evidence or without an adequate account of its reasoning, then the Appellate Body should not rule that the panel had made an "objective assessment of the matter before it." See *id.* Therefore, this standard of reasonableness derives from the DSU's prescription of objectivity.

179. On the concept of minimally sufficient evidence, see generally JAMES ET AL., *supra* note 169, §§ 7.19, 12.9; MCCORMICK ON EVIDENCE, *supra* note 169, § 338. But see KAZAZI, *supra* note 169, at 21-38, 368 ("[T]here is no place in international proceedings for the Anglo-American dual concept of the burden of proof," which distinguishes burden of persuasion, on the one hand, and burden of production or sufficiency of evidence, on the other.). However, the reasons given by Kazazi — that there are no technical rules of evidence and no jury in international proceedings — are not persuasive in the WTO context. What drives the development of a concept of minimally sufficient evidence and at least some rules of evidence is not merely the presence of a jury, but also the existence of an appellate body that may set aside only those findings of fact that are erroneous as a matter of law.

example, if a panel should have taken into account highly relevant and available evidence but failed to do so, then it has acted in an arbitrary and unreasonable manner.¹⁸⁰ Upon appellate review, the Appellate Body should determine whether a panel's record contained the minimal rational support needed for the panel's findings.¹⁸¹ A panel could also act unreasonably by failing to give any account of its reasoning from the evidence.¹⁸² The scope of the Appellate Body's review of a panel's findings of fact should be a deferential one. So long as a panel acts reasonably in basing its findings on the evidence before it, the Appellate Body should not substitute its judgment for that of the panel. If, however, a panel acts arbitrarily in making its findings, thereby violating the minimum requirements of rational inference, then its findings should be set aside.¹⁸³

Finally, with respect to any particular factual issue, the parties may

180. Cf. *Citizens to Preserve Overton Park, Inc. v. Volpe*, 401 U.S. 402, 416, 91 S. Ct. 814, 823-24 (1971) (holding that in order to make a finding that agency action was not "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law," an American reviewing court "must consider whether the [agency's] decision was based on a consideration of the relevant factors and whether there has been a clear error of judgment").

181. Article 12.7 of the DSU provides that "the report of a panel shall set out the findings of fact, the applicability of relevant provisions and the basic rationale behind any findings and recommendations that it makes." DSU, *supra* note 7, art. 12.7. Thus, the DSU requires panels to provide in their reports the information needed for meaningful Appellate Body review. See *infra* note 183.

182. See DSU, *supra* note 7, art. 12.7 and *infra* note 183; cf. *Overton Park*, 401 U.S. at 415-21, 91 S. Ct. at 823-26 (remanding case where reviewing district court had failed to create a judicial record that was adequate to support the findings required in judicial review).

183. The Appellate Body's review of a panel's findings about scientific plausibility should be a very different process than a panel's review of a member's scientific determination. The lack of parallel derives from both the record on review and the issues on review. As for the record on review, the Appellate Body should consider only evidence that was in the record before the panel. However, the panel has the authority and should have the obligation to create a record containing current scientific information and should not restrict the scientific evidence it considers to evidence from a certain time period or domestic proceeding in the past. See *supra* Part II.B. Both the Appellate Body and a panel must draft their reports "in the light of the information provided and the statements made" before them. DSU, *supra* note 7, arts. 17.10, 14.2. Unlike the Appellate Body, however, a panel has "the right to seek information and technical advice from any individual or body which it deems appropriate." *Id.* art. 13; cf. *id.* app. 3.8. Moreover, if a panel requests an advisory report from an "expert review group" in accordance with article 13.2, that expert review group in turn has the authority to "consult and seek information and technical advice from any source they deem appropriate." *Id.* app. 4.4. Additionally, any member is required to "respond promptly and fully to any request by an expert review group for such information as the expert review group considers necessary and appropriate." *Id.* Unlike the Appellate Body, a panel has the mandate and authority to gather that evidence which it considers relevant to its factfinding.

As for the issues on review, the Appellate Body should determine whether the panel's record contained an adequate basis for the findings the panel made. For example, the Appellate Body should determine whether the panel's record contained the minimum evidence necessary to sustain a finding of scientific plausibility. The factfinding panel, by contrast, should use its best judgment to find whether the scientific evidence produced in its own proceeding proves, by a preponderance, that a member's measures have reasonable scientific support. But so long as there exists minimally sufficient scientific evidence in the panel's record and a reasonable person could have arrived at the panel's

have various burdens of producing evidence placed on them.¹⁸⁴ A panel is not primarily an investigative institution with the capacity to generate all of its own evidence. Consequently, a panel must have rules that dictate when a party must produce evidence to the panel, and how much evidence that party must produce in order to avoid summary decision against it. Under the SPS Agreement, a complaining party "must establish a *prima facie* case of inconsistency with a particular provision of the SPS Agreement . . ."¹⁸⁵ A member claiming a violation of a fundamental obligation under an agreement "must assert and prove its claim."¹⁸⁶ That is, the complaining party must produce a *prima facie* case on any factual issues essential to its claim.¹⁸⁷ The phrase "*prima facie* case" may have different meanings in different legal settings.¹⁸⁸ For purposes of WTO factfinding under the SPS Agreement, evidence should constitute a *prima facie* case if (a) it is sufficient to provide the minimum rational support for an inference to the requested finding and (b) it would be in fact persuasive to the panel if no contrary evidence were produced. A *prima facie* case must therefore be of sufficient quality to pass appellate review for reasonableness and of sufficient weight to persuade the panel in the absence of rebuttal evidence.

conclusion, then the Appellate Body must defer and not substitute its judgment for that of the panel.

184. On the burden of production, see JAMES ET AL., *supra* note 169, § 7.15; MCCORMICK ON EVIDENCE, *supra* note 169, § 338.

185. Appellate Body Report, *supra* note 16, ¶ 98; see Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 13. Although several GATT and WTO panels have placed on defending members the burden of proving an exception under GATT Article XX, the Appellate Body has held that these provisions are "limited exceptions from obligations under certain other provisions of the GATT 1994, not positive rules establishing obligations in themselves." *Id.* at 16. The burden of establishing such "affirmative defenses" should rest on the party asserting them. When, however, a complaining party alleges a violation of a provision setting down "a fundamental part of the rights and obligations of WTO Members," that party "must assert and prove its claim." *Id.*

The Hormones Panel reasoned that the fact that article 3.2 of the SPS Agreement explicitly establishes a presumption that a *conforming* measure is consistent with both the SPS Agreement and GATT 1994 suggests that the "burden of proof" in all other situations rests on the defending member, or at least "shifts" the burden to that member after some initial showing by the complaining member. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.57. This reasoning was rejected by the Appellate Body as a misinterpretation of the SPS Agreement. Appellate Body Report, *supra* note 16, ¶¶ 102-05.

186. See Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 16.

187. For a discussion of the phrase "*prima facie* evidence" in the context of international tribunals, see KAZAZI, *supra* note 169, at 326-43 (discussing meaning of "*prima facie* evidence" and reporting the "generally uncontested" definition that it is evidence which, if unexplained or uncontradicted, is sufficient to maintain the proposition affirmed; such evidence "shifts the burden of evidence from the proponent of the burden of proof to the other party," where "burden of evidence" means "burden of presenting evidence," not "burden of proof").

188. See *id.* In another context, the DSU refers to an infringement of the obligations assumed under a covered agreement as "*prima facie* to constitute a case of nullification or impairment." DSU, *supra* note 7, art. 3.8. In article 3.8, "*prima facie* case" means that "there is normally a presumption" that a *breach* of the rules has an *adverse impact* on other members, and that "it shall be up to the Member against whom the complaint has been brought to rebut the charge." *Id.* This DSU provision is consistent with the interpretation of "*prima facie*" proposed in this Article.

By definition, therefore, once a complaining party has produced a *prima facie* case, the member against whom the complaint has been brought faces a tactical decision: either produce convincing rebuttal evidence or suffer adverse findings.¹⁸⁹ It may be misleading, however, to speak of any "shift" in the "burden of proof" in such a situation.¹⁹⁰ The burden of persuasion on any issue essential to proving infringement of the Agreement should not be placed on the defending party. Moreover, a formal burden of production should be assigned to a defending party only in very unusual circumstances, such as when that party has peculiar access to important evidence.¹⁹¹ In the dynamics of the normal proof process, there may come a point when the defending party will lose unless it produces convincing rebuttal evidence. But this does not mean that any burdens of persuasion or production have been shifted to that party.

This allocation of roles in the factfinding process is consistent with the general presumption of good-faith compliance for members, as well as with the substantive provisions of the SPS Agreement.¹⁹² In disputes involving the SPS Agreement, the complaining member should produce a *prima facie* case that the Agreement has been violated. In disputes involving article 2.2, for example, a complaining member should have both the burden of production and the burden of persuasion in proving that there is no reasonable scientific basis for the adopted sanitary measure.¹⁹³ Under article 2.2, the complaining member should also have the burden of proving that the challenged measure is not necessary to achieve the defendant member's selected level of protection.¹⁹⁴ In complaints alleging a violation of article 3, the complaining member should have the burden to prove which of the three options provided in articles 3.1, 3.2, and 3.3 applies and has been violated.¹⁹⁵ This allocation of burdens is consistent not only with the gen-

189. See Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 13.

190. For example, the report of the Appellate Body in the woven wool shirts case stated, somewhat ambiguously:

[I]t is a generally-accepted canon of evidence . . . that the burden of proof rests upon the party . . . who asserts the affirmative of a particular claim or defence. If that party adduces evidence to raise a presumption that what is claimed is true, the burden then shifts to the other party, who will fail unless it adduces sufficient evidence to rebut the presumption.

Id. What is unclear from this decision is whether the burden that "shifts" is merely a burden to produce additional evidence (and if so how much is needed to "rebut a presumption"), or whether the burden of persuasion also "shifts" to the defending party (who now must lose if the total evidence is in "equipoise").

191. Cf. KAZAZI, *supra* note 169, at 119-51, 223, 357-59 (discussing duty of collaboration of parties to international proceedings, including the duty to provide explanations and documents that may be in a party's sole possession).

192. As the Appellate Body held in the context of the Agreement on Textiles and Clothing, when interpreting "'carefully negotiated language . . . which reflects an equally carefully drawn balance of rights and obligations of Members . . .,' [t]hat balance must be respected." Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 16 (quoting from Report of the Appellate Body, United States — Restrictions on Imports of Cotton and Man-made Fibre Underwear, Feb. 25, 1997, WT/DS24/AB/R, at 15).

193. See SPS Agreement, *supra* note 3, art. 2.2; *supra* Parts I.A, I.B, II.A.

194. See SPS Agreement, *supra* note 3, art. 2.2; *supra* Parts I.B.2, II.A.2.

195. See SPS Agreement, *supra* note 3, arts. 3.1-3.3; *supra* Parts I.C, II.A.2.

eral principles discussed above, but also with other SPS provisions.¹⁹⁶

The roles and responsibilities in the factfinding process that are proposed in this part of the Article would bring about the fair and efficient resolution of disputes arising under the SPS Agreement and would respect the sovereign right of members to protect health and life within their territories. What remains to be examined is how close to this model the WTO dispute settlement institutions come in practice. For this purpose, the next part examines the WTO's factfinding process in an actual dispute.

III. Factfinding in the Growth Hormones Proceeding

Inherent in the reports of the WTO Hormones Panel¹⁹⁷ and in the Appellate Body Report¹⁹⁸ are the elements of a general approach to future SPS disputes. The aspect of this approach relevant to this Article is the Panel's and Appellate Body's treatment of the issues of scientific uncertainty, science policy, and burden of proof. In what follows, I examine their treatment of these issues and evaluate their general approach to SPS dispute settlement. I conclude that, on a number of important points, when compared to the analysis presented earlier in this Article, the approach taken by the Panel and the Appellate Body remains unclear or is incorrect.

A. The Proper Substantive Issues Before the Hormones Panel

The first point of comparison is the Panel's and Appellate Body's understanding of the content of the findings to be made. The content varies, depending upon whether the Panel is making findings with respect to risk assessment or to risk management. My affirmative analyses of risk assessment and risk management, as well as of the content of WTO findings concerning them, are set forth above, in Parts I and II.A, respectively.

1. Content of Findings on Risk Assessment Determinations

The definition of "risk assessment" employed by the Hormones Panel closely follows the conception set forth in the SPS Agreement.¹⁹⁹ Moreover, the Panel found no techniques developed by relevant international

196. SPS article 4.1 expressly places on the exporting (complaining) member the burden of proving that an alternative sanitary measure is as protective as that chosen by the importing member. See SPS Agreement, *supra* note 3, art. 4.1. Under SPS article 5.6, the exporting (complaining) member should have the burden to prove that there exists an alternative sanitary measure that would achieve the same level of protection selected by an importing member and that would be significantly less restrictive to trade. See *id.* art. 5.6.

197. Hormones Panel Report (US), *supra* note 13; Hormones Panel Report (CAN), *supra* note 13.

198. Appellate Body Report, *supra* note 16.

199. See SPS Agreement, *supra* note 3, art. 5, annex A.4; Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.95, 8.101 (concluding that a risk assessment for purposes of the SPS Agreement should "identify the adverse effects on human health" and, if any such adverse effects exist, should "evaluate the potential or probability of occurrence of these effects"). The Appellate Body Report expressed a strong preference for not departing from the sparse language of the SPS Agreement, although it labelled as not "substantially wrong" the Panel's elaboration of risk assessment as a two-step process that first identifies adverse effects and then identifies their potential for occurrence. Appellate Body Report, *supra* note 16, ¶¶ 182-84.

organizations that had to be taken into account in risk assessment.²⁰⁰ Presumably, therefore, members were free to develop and apply risk assessment techniques "as appropriate to the circumstances," so long as the challenged sanitary measure was "based on an assessment . . . of the risks" and that assessment of risks incorporated "scientific principles" and contained "sufficient scientific evidence."²⁰¹

The Hormones Panel began by asserting that "it is not for the Panel itself to conduct its own risk assessment on the basis of scientific evidence."²⁰² This assertion raises the question of what exactly a panel is supposed to do with respect to risk assessment. I argued above that the Hormones Panel should have made findings about whether there existed, at the time of the Panel's proceeding, any reasonable scientific basis for the European Communities' challenged sanitary measures, and whether the EC's science-policy choices employed scientifically plausible alternatives.²⁰³ However, the Hormones Panel did not follow this approach. Instead, it developed a two-part, procedural and substantive interpretation of the "based on" language of SPS article 5.1. This approach led to a number of serious errors.

For the first part of its interpretation, the Panel found that the requirement of SPS article 5.1, that a sanitary measure be "based on an assessment . . . of the risks," included a "minimum procedural requirement."²⁰⁴ This requirement was that "the Member imposing a sanitary measure needs to submit evidence that at least it actually *took into account* a risk assessment when it enacted or maintained its sanitary measure in order for that measure to be considered as *based on* a risk assessment."²⁰⁵ For the Panel, this requirement meant that the EC had to produce documentation of a past risk assessment that it had considered when adopting or maintaining its measures.²⁰⁶ Therefore, the Panel required the production of

200. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.106.

201. See SPS Agreement, *supra* note 3, arts. 5.1, 2.2; see *supra* Part I.A. The Appellate Body Report "stress[ed] that Articles 2.2 and 5.1 should constantly be read together. Article 2.2 informs Article 5.1" Appellate Body Report, *supra* note 16, ¶ 180.

202. Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.104, 8.118.

203. See *supra* Parts I.A, II.A.

204. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.116; SPS Agreement, *supra* note 3, art. 5.1.

205. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.116; see *id.* ¶ 8.162; SPS Agreement, *supra* note 3, art. 5.1.

206. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.114, 8.111-8.137. The Panel did acknowledge that, in the case of "a sanitary measure enacted *before* the entry into force of the SPS Agreement," the risk assessment itself might not come into existence until *after* the sanitary measure. *Id.* ¶ 8.102. In such a case, the risk assessment would show that the measure was maintained in compliance with the Agreement.

The Panel routinely used the phrase "a risk assessment," which provides grammatical encouragement for reifying the risk assessment as a physical document. This noun phrase introduced by an indefinite article, however, is not used in the SPS Agreement. The closest counterpart appears to be in article 5.1, which refers to "an assessment, as appropriate to the circumstances, of the risks." SPS Agreement, *supra* note 3, art. 5.1. But this latter phrasing could easily refer to a line of reasoning or a mode of justification, instead of a document. This would be in keeping with the phrasing of SPS annex A.4, which defines "risk assessment" as "the evaluation of" the potential for adverse effects.

one or more risk assessment documents created prior to the panel proceeding.²⁰⁷

The Appellate Body held that the Panel had created a new procedural obligation that is not in the SPS Agreement.²⁰⁸ The Appellate Body held that a member is not required to conduct its own risk assessment, but could defend its measures by relying on a risk assessment conducted by another member or by an international organization.²⁰⁹ The Appellate Body noted, consistent with the argument above, that it would be unwise and unwarranted to disregard available scientific evidence simply because it had not been included by the defending member in a prior risk assessment.²¹⁰

In another respect, however, the Appellate Body apparently agreed with the Hormones Panel that members maintaining sanitary measures must be able to produce for a panel already completed studies that do constitute an adequate risk assessment. For example, in the case of the hormone MGA, there existed little or no public information at the time of the Hormones Panel Report.²¹¹ Nevertheless, the Appellate Body left undisturbed the Panel's finding that there was no risk assessment for MGA, and held that the European Communities had to produce such a risk assessment.²¹² Thus, although the Appellate Body did not require historical evi-

SPS Agreement, *supra* note 3, annex A.4. A line of reasoning or mode of justification is not identical to a physical document. Another possibility is that "risk assessment" is an ongoing evaluative process. This sense of the term is supported by the grammatical phrasing of SPS article 5.2 ("In the assessment of risks") and of article 5.3 ("In assessing the risk"). *Id.* arts. 5.2, 5.3. Further, annex A.1 to the SPS Agreement lists "methods of risk assessment" as an example of a *sanitary measure* that a member might establish, and annex B.3(c) to the SPS Agreement requires that "one enquiry point" on "risk assessment procedures" be established by the member — because a member's risk assessment *procedures* could themselves be regulatory requirements that are restrictive of trade. *Id.* annexes A.1, B.3(c). The SPS Agreement therefore suggests by its language that "a risk assessment" as a document or procedure should be distinguished from "an assessment of risks" in the sense of a line of reasoning or a mode of justification. The Hormones Panel disregarded this linguistic distinction by concluding that the SPS Agreement requires members to be able to produce a prior document called "a risk assessment." See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.111-8.114.

207. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.116-8.119. Although the Panel occasionally used ambiguous language, its intent seems very clear. The Panel found that the so-called "new evidence" gathered by the European Communities during the Panel process can, therefore, from a procedural point of view, not be considered as part of a risk assessment on which the European Communities based its measures in dispute." *Id.* ¶ 8.118.

208. Appellate Body Report, *supra* note 16, ¶ 189.

209. *Id.* ¶ 190 (stating that a member adopting a sanitary measure need not "have carried out its own risk assessment," but the measure could have "its objective justification in a risk assessment carried out by another Member, or an international organization").

210. See *id.* ¶¶ 189-90; *supra* Part II.B.

211. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.258, 8.264; Appellate Body Report, *supra* note 16, ¶ 201.

212. The Appellate Body Report noted that "there was an almost complete absence of evidence on MGA in the panel proceedings. We therefore uphold the Panel's finding that there was no risk assessment with regard to MGA." Appellate Body Report, *supra* note 16, ¶¶ 201, 208.

dence of the EC's past reliance on risk assessment studies, it did require that such studies exist and be produced in a dispute settlement proceeding. The Appellate Body therefore approved one aspect of the Panel's procedural requirement: the defending member must have performed or obtained adequate risk assessment studies prior to the panel's request for them.

In addition, those studies had to be, at least in the hormones dispute, *specifically* focused on "the carcinogenic or genotoxic potential of the residues of those hormones found in meat derived from cattle to which the hormones had been administered for growth promotion purposes."²¹³ It might be argued that such specificity is not required as a general or procedural matter, but only because the affirmative evidence in the hormones dispute weighed so overwhelmingly against the EC's position.²¹⁴ Such an argument is undermined, however, by the sole paragraph devoted to the adequacy of risk assessment for MGA under article 5.1.²¹⁵ Although "there was an almost complete absence of evidence on MGA in the panel proceedings,"²¹⁶ the studies in the record addressing the carcinogenic potential of progestins²¹⁷ were held not to be a risk assessment for MGA because the EC had not produced "any study that demonstrated how closely related MGA is chemically and pharmacologically to other progestins and what effects MGA residues would actually have on human beings when such residues are ingested along with meat from cattle to which MGA has been administered for growth promotion purposes."²¹⁸ In view of the Appellate Body's treatment of MGA, it is reasonable to conclude that members who wish to establish or maintain sanitary measures are now obliged to provide scientific studies that assess the risk posed by specific chemical

213. *Id.* ¶¶ 198-200. According to the Appellate Body, such specificity is required by annex A.4 to the SPS Agreement (providing the definition of risk assessment). *Id.* ¶ 200.

214. *See id.* ¶¶ 192-209. For example, the Appellate Body raised the issue of sufficient specificity when it evaluated the EC's rebuttal evidence concerning risk. *Id.* ¶ 198 (finding that the opinion by Dr. Lucier "does not purport to be the result of *scientific studies* carried out by him or under his supervision *focusing specifically* on residues of hormones in meat from cattle fattened with such hormones" and thus "is not reasonably sufficient to overturn the contrary conclusions reached in the scientific studies . . . that related *specifically* to residues of the hormones in meat from cattle to which hormones had been administered for growth promotion") (emphasis added).

215. *Id.* ¶ 201.

216. *Id.*

217. There is evidence in the record that progestins, including progesterone, are possibly carcinogenic, Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.129-8.131, and that the synthetic hormone MGA mimics the biological activity of progesterone. *Id.* ¶ 2.8. *See id.* ¶¶ 4.151-4.171.

218. Appellate Body Report, *supra* note 16, ¶ 201. This conclusion was consistent with the general finding of the Appellate Body that:

[t]he 1987 IARC Monographs and the articles and opinions of individual scientists submitted by the European Communities constitute general studies which do indeed show the existence of a general risk of cancer; but they do not focus on and do not address the particular kind of risk here at stake — the carcinogenic or genotoxic potential of the residues of those hormones found in meat derived from cattle to which the hormones had been administered for growth promotion purposes. . . .

agents when used in specific ways.²¹⁹

Even the Appellate Body's procedural requirement, however, is an unreasonable interpretation of the SPS Agreement. Such an interpretation implies that every sanitary measure existing at the time the SPS Agreement entered into force suddenly became inconsistent with the Agreement unless and until someone completed specifically focused studies justifying the measure. This obligation implies that it is inconsistent with the SPS Agreement to maintain a domestic regulatory program that automatically bars import of a new product until the product supplier produces sufficient data and proves its safety. Such pre-market approval programs impose bans on products for which there are no specific data, studies, or formal risk assessment.²²⁰

However, the SPS Agreement should not be read to impose on importing members the burden of having complete and specific studies for every trade-restrictive effect of a sanitary measure.²²¹ Article 5.1 merely requires that the measure be "based on an assessment, as appropriate to the circumstances, of the risks."²²² Pre-market approval programs for untested pesticides or untested food additives can surely be risk-based measures even in the absence of substance-specific studies.²²³ What makes such programs risk-based is the fact that they are designed to address broad categories of risk: using chemicals that are toxic by design (pesticides) and eating food additives that have unknown characteristics. Moreover, in the case of MGA, there is affirmative evidence that MGA mimics a compound with carcinogenic potential (progesterone).²²⁴ In cases where such affirmative evidence exists, the question for the panel should be whether this evidence constitutes a plausible basis for the conclusion that there is a risk. Thus, article 5.1 should not be interpreted as creating any procedural requirements, even a requirement to conduct or locate risk assessment studies. The only requirement should be that the challenged measure satisfies the substantive SPS standards for risk assessment, to which we now turn.

In the second and substantive part of its interpretation, the Panel considered its task to be one of comparing the scientific conclusions reached

Id. ¶ 200.

219. The Appellate Body grounded this obligation on the SPS Agreement's definition of "risk assessment" as "the evaluation of the potential for adverse effects on human or animal health arising from the presence of additives, contaminants, toxins or disease-causing organisms in food, beverages or feedstuffs." *Id.* ¶ 200 (citing the SPS Agreement, *supra* note 3, annex A.4).

220. See, e.g., 7 U.S.C. § 136a (1994) and 40 C.F.R. Part 158 (1997) (providing for pesticide registration); 21 U.S.C. §§ 331, 342(a)(1)(C), 348 (1994) and 21 C.F.R. Part 171 (1997) (providing for regulations allowing use of food additives).

221. If the SPS Agreement makes too little provision for a transition period in which resource-intensive risk assessments could be conducted in a deliberate fashion on a priority basis without disrupting existing sanitary measures or trade patterns, that problem should not be exacerbated by holding that the assessment of risks must in all cases be documented by a report completed prior to the WTO panel proceeding.

222. SPS Agreement, *supra* note 3, art. 5.1.

223. See *supra* note 220.

224. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 2.9, 4.154, 4.161-4.162, 4.164-4.167, 8.129, n.440, 8.131.

in the relevant risk assessment documents with the scientific conclusions being asserted by the European Communities in the panel proceeding. The Panel's goal was to determine whether the EC's assertions were "in conformity with any of those reached in the [risk assessment] studies."²²⁵ When the Panel found that the EC's conclusion on the safety of the growth-promoting use of hormones did "not conform to any of the scientific conclusions reached" in the studies produced in the Panel proceeding, the Panel concluded that the EC's sanitary measures were "not based on the scientific evidence submitted to the Panel."²²⁶ There are, however, several serious problems with this approach.

The most obvious difficulty is that it turns the scientists who conduct risk assessments into the *de facto* factfinders for the WTO. Consequently, neither members nor WTO panels would be able to make independent assessments on the merits. Under the Panel's approach, a member's sanitary measures would satisfy article 5.1 only if the member had empaneled or found a group of scientists who had stated a conclusion consistent with the position of the member. Moreover, this approach could easily devolve into a mere procedural formality, with members convening formal meetings of only sympathetic scientists. A panel should instead determine independently whether there is any reasonable scientific basis for a member's risk determinations and sanitary measures.²²⁷ Ceding the factfinding authority of the WTO to past convocations of scientists is a development that is both inappropriate and too easy to circumvent.²²⁸

Moreover, the Panel's interpretation understandably led it to inconsistent conclusions. The Panel appointed six experts to advise it on scientific matters.²²⁹ Although the Panel asserted that it was not its task to conduct its own risk assessment,²³⁰ it turned to its expert advisors to help it to decide both the meaning and the merits of the risk assessment documents it had before it.²³¹ Moreover, the Panel itself decided the truth of various

225. *Id.* ¶ 8.120; Appellate Body Report, *supra* note 16, ¶ 192.

226. Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.140, 8.162.

227. *See supra* Parts I.A., II.A.

228. Moreover, the Panel's approach would create insoluble and unnecessary questions about what is to count as a qualifying scientific report. The Hormones Panel dodged the difficulties in this issue by "assuming" that the documents referred to by the European Communities met the "minimum requirements of a risk assessment." Hormones Panel Report (CAN), *supra* note 13, ¶ 8.114. In future disputes the parties might not be so reasonable. Is a minority report from a scientific conference "a risk assessment"? Is a single co-authored study published in a peer-reviewed scientific journal? Are there to be minimum WTO requirements for a bona fide "risk assessment process"? One difficulty with a purely procedural requirement is that we would almost certainly need substantive criteria for applying it. For example, the Hormones Panel seemed to be uncertain about what to do with studies that appear to be *generally* relevant to hormones but which do not discuss the *specific* uses at issue in the dispute. *See id.* ¶¶ 8.130, 8.133.

229. *See id.* ¶¶ 6.9, 8.7-8.9.

230. *See id.* ¶¶ 8.104, 8.118.

231. *See id.* ¶ 8.127. The Hormones Panel stated that:

[a]ll of the scientific studies outlined above came to the conclusion that the use of the hormones at issue (all but MGA, for which no evidence was submitted)

scientific propositions relevant to the case.²³² The reason for this tension seems clear. A panel cannot merely *look* at a document reporting the results of a scientific conference to see whether the report calls itself a "risk assessment" and what words are used in stating the report's conclusions. To make its decisions, a panel has to *understand* the meaning of the assertions made by scientists. The Hormones Panel took the worst approach: (1) pretend not to pass judgment on the merits of past scientific reports²³³ while implicitly agreeing with their conclusions, and (2) formally preclude rebuttal explanation in the WTO proceeding by members who disagree with those scientific reports, saying that WTO panels do not themselves conduct risk assessments.

The Appellate Body refocused attention on the need to evaluate the meaning and content of scientific evidence. The Appellate Body held that the Panel's approach of matching the member's arguments against the conclusions found in a risk assessment document might be useful and relevant, but that this was not sufficient to exhaust the meaning of "based on" a risk assessment.²³⁴ More is required of a panel than merely matching documents. In developing its own interpretation of "based on," the Appel-

for growth promotion purposes is safe; . . . We note that this conclusion has also been confirmed by the scientific experts advising the Panel.

See *id.* ¶ 8.136 ("[T]he scientific experts advising the Panel were of the view that this evidence, as it stands today, does not invalidate or contradict the scientific conclusions reached in the other studies invoked by the European Communities . . ."); *id.* ¶ 8.137 ("[A]ll scientific experts advising the Panel confirmed this conclusion and stated that, as of today, no scientific evidence is available which concludes that an identifiable risk arises from the use of any of the hormones at issue for growth promotion purposes in accordance with good practice."); *id.* ¶ 8.114 ("[T]he scientists advising the Panel seemed to consider these reports, from a scientific and technical point of view, to be risk assessments.").

232. See *id.* ¶ 8.137 ("we find that the European Communities has not demonstrated that the scientific evidence it referred to . . . would indicate that an identifiable risk arises for human health"); *id.* ¶¶ 8.131-8.132 ("we consider that the 1987 IARC Monographs . . . do not contradict, the other studies referred to by the European Communities . . .").

233. The Hormones Panel went to great lengths to appear not to make any scientific determinations. At one point in its deliberations about the relevance and significance of several articles and opinions of individual scientists that "deal with the carcinogenic or genotoxic potential of hormones and criticize the scientific methodology or conclusions of the other studies referred to by the European Communities," *id.* ¶ 8.133, the Panel rather implausibly concluded:

according to the Codex expert advising the Panel, most of the evidence contained in these articles and opinions and the potential risks addressed therein were *already evaluated and taken into account* in the 1988 and 1989 JECFA Reports. Indeed, in the event these articles and opinions should be considered as evidence which was . . . not "new" but was already taken into account in the 1988 or 1989 JECFA Reports, the Lamming Report or the 1995 EC Scientific Conference, these articles and opinions *would not then invalidate or contradict the scientific conclusions reached in these other studies* . . .

Id. ¶ 8.135 (emphasis added). The Panel retreated to this implausible position (that if a study was taken into account at a conference, then its results could not be inconsistent with the findings of the conference) presumably because it was trying to avoid the need to understand the issues and arguments themselves, and was trying to deal with the documents only on their face, devoid of meaning.

234. See Appellate Body Report, *supra* note 16, ¶ 193.

late Body arrived at its own statement of the content of a panel's findings on risk assessment:

We believe that Article 5.1 . . . requires that the results of *the risk assessment* must *sufficiently warrant* — that is to say, *reasonably support* — the SPS measure at stake. The requirement that an SPS measure be "based on" a risk assessment is a substantive requirement that there be a *rational relationship* between the measure and *the risk assessment*.²³⁵

This search for "reasonable support" implies that a risk assessment document need not have reached a "monolithic conclusion" that supports the member's position; a member might "act in good faith on the basis of what, at a given time, may be a divergent opinion coming from qualified and respected sources."²³⁶ What is required is that there be some available scientific evidence that qualifies as "a risk assessment" and which provides "sufficient warrant" or "reasonable support" for a member's sanitary measure.

While the Appellate Body's formulation is much closer to my proposal than is the Panel's formulation, the Appellate Body's position might still fall short, depending upon how it is interpreted in future cases. First, the Appellate Body may require a measure to be reasonably supported by a risk assessment that is so "specifically focused" that the risk determination itself must clear a high threshold of specificity.²³⁷ In addition, the Appellate Body has not sufficiently clarified the roles of scientific uncertainty and science policy in risk assessment. When there are two plausible scientific accounts and it is uncertain which will prove to be correct, then a member should be entitled, without further explanation, to base its regulation on either account, as a matter of its own science policy.²³⁸ I have argued that a member's right to set its own science policy within risk assessment is guaranteed by the SPS Agreement.²³⁹ In its careful parsing of the SPS Agreement, the Appellate Body overturned some of the more glaring errors of this particular factfinding panel, but it did not set out sufficient guidelines on how this issue should be addressed by future panels.

The Appellate Body's inadequate attention to scientific uncertainty and science policy is illustrated in its treatment of the concepts of "safety" and "risk." The Hormones Panel preferred to understand "safety" as an all-

235. *Id.* (emphasis added).

236. *Id.* ¶ 194.

237. See *supra* text accompanying notes 213-20.

238. I have argued that, as a matter of law, the existence of a good-faith controversy among qualified scientists should constitute compelling evidence of scientific uncertainty and of the scientific plausibility of the competing views. See *supra* Parts II.A, II.C. The Appellate Body Report in the hormones dispute acknowledged that such a controversy may indicate or be "a form of" scientific uncertainty, but concluded only that a member's good-faith reliance on one side of the controversy "does not necessarily signal the absence of a reasonable relationship between the SPS measure and the risk assessment." Appellate Body Report, *supra* note 16, ¶ 194. The Appellate Body refrained, however, from holding that such a controversy should be compelling evidence of uncertainty and plausibility. See *id.*

239. See *supra* Part II.A.

or-nothing *scientific* issue.²⁴⁰ But most risk assessments today, due to the presence of scientific uncertainty, are not purely scientific. Risk assessors must employ science policy in order to complete risk assessments. As argued above,²⁴¹ the assessment of risk with respect to carcinogens normally involves the pervasive use of science policies to resolve scientific uncertainties, and those policies reflect risk management policies on what risk is "safe" or acceptable.²⁴² The Appellate Body found that the Hormones Panel had no grounds in the SPS Agreement for distinguishing between risk assessment and risk management, and for relegating policy decisions exclusively to the latter.²⁴³ In keeping with this finding, the Appellate Body virtually ignored risk management because the phrase is not mentioned in the SPS Agreement.²⁴⁴ Thus, the Appellate Body perhaps left room to locate science policy within the concept of risk assessment, but it did so in an entirely negative way. This negative approach ignores a member's sovereign right to establish its own science policy, which derives its content and rationale from risk management and from the right to select an appropriate level of protection.²⁴⁵ At a minimum, a panel should identify and defer to those science-policy choices of the defending member that result in a scientifically plausible assessment of risk. In most cases, panels will find that determinations of "safety" are not purely scientific, but are also based on assumptions made pursuant to science policies.

The Appellate Body also dealt inadequately with the concept of risk. The Appellate Body corrected the mistaken view of the Hormones Panel that a risk must be identified or characterized quantitatively by assigning a probability to the likelihood of occurrence.²⁴⁶ In addition, the Appellate Body was rightly concerned that, by substituting the word "probability" for the word "potential" in the SPS Agreement's definition of "risk assessment,"²⁴⁷ the Panel introduced "a minimum magnitude of risk" as a *de minimis* threshold of concern under the SPS Agreement, when the Agree-

240. The Panel stated:

In our view, the scientific conclusion reflected in the EC measures in dispute, i.e., that the use of the hormones in dispute for growth promotion purposes, even in accordance with good practice, is *not* safe, does not conform to any of the scientific conclusions reached in the evidence referred to by the European Communities. All the evidence referred to by the European Communities which specifically relates to the use of the hormones at issue for growth promotion purposes concludes that the use of these hormones as growth promoters in accordance with good practice is safe.

Hormones Panel Report (CAN), *supra* note 13, ¶ 8.140.

241. See *supra* Parts I.A, I.B.

242. This position was apparently taken by the European Communities. See Hormones Panel Report (CAN), *supra* note 13, ¶ 8.152.

243. See Appellate Body Report, *supra* note 16, ¶ 181 ("Thus, the Panel's distinction [between risk assessment and risk management], which it apparently employs to achieve or support what appears to be a restrictive notion of risk assessment, has no textual basis [in the SPS Agreement].").

244. *Id.* ¶¶ 181, 206.

245. See *supra* Parts I.A, I.B, II.A.

246. Appellate Body Report, *supra* note 16, ¶ 184.

247. *Id.*; SPS Agreement, *supra* note 3, annex A.4.

ment provides no basis for such a threshold.²⁴⁸ On the other hand, the Appellate Body found that the kind of risk arising from a merely "theoretical uncertainty" is not adequate to justify protective measures under the Agreement.²⁴⁹ Instead, the Appellate Body concluded that there must be a justification based on an "ascertainable risk," although the methods of demonstrating its existence need not be "quantitative analysis by the empirical or experimental laboratory methods commonly associated with the physical sciences."²⁵⁰ But the Appellate Body Report failed to pursue its own reasoning on this issue far enough.

What is needed is an adequate account of the nature of risk, which includes an understanding of how scientific uncertainty can itself reflect risk. The problem is not merely that the phrase "a probability" connotes a *higher* degree of likelihood as compared to "a potential" or "a possibility," but that the phrase can focus our attention on quantification alone. Risk is measured not only by positive knowledge of a quantifiable likelihood, but also by the degree of uncertainty or lack of knowledge about a possible hazard. Choices made in the face of uncertainty between two plausible alternatives reflect a true risk. Indeed, the risk taken in acting on such a choice is directly proportional to the chance of error inherent in any underlying prediction. On the continuum between a merely speculative risk and a conclusively demonstrated one lies a vast stretch of undemonstrated, unquantified, but scientifically plausible risks. Within that zone, the risk of harm is real so long as safety is unproven.

A member's right under the SPS Agreement to adopt any level of protection is also the right to characterize or assess as a real risk any adverse effect that is "possible," in the sense of "scientifically plausible."²⁵¹ The SPS Agreement guarantees the member's right to adopt science policies for conducting risk assessment in the face of scientific uncertainty, to find that a scientifically plausible set of assumptions does identify a real risk, and to choose to protect those individuals within its territory from incurring that risk. By failing to take into account the important role of science policy in assessing risk, the Appellate Body also failed to address the right of members to find true risks wherever adverse effects are scientifically plausible. A WTO panel should have no discretion to find a member's measure *not* "sufficiently warranted" or *not* "reasonably supported" by the scientific evidence, if the panel finds that there is at least one scientifically plausible set of assumptions under which an adverse effect might occur.

2. Content of Findings on Risk Management Determinations

The Hormones Panel apparently began with the same conception of risk management set forth earlier in this Article.²⁵² The Panel recognized that

248. Appellate Body Report, *supra* note 16, ¶¶ 184-86.

249. *See id.* ¶ 186.

250. *See id.* ¶¶ 186-87.

251. *See supra* Parts I.A, I.B, II.A.

252. Compare Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.98, 8.163, with *supra* Part I.B.

risk management involves decisions about risk acceptance that are political and based on non-scientific factors.²⁵³ By insisting that risk assessment is scientific and risk management is value-based,²⁵⁴ the Panel rejected the possibility that risk management might set science policies for conducting risk assessments. The Panel also tended to underestimate the role of science in the cost-benefit analyses inherent in management decisions.²⁵⁵ As pointed out in the previous part, the Appellate Body virtually ignored the concept of "risk management" because it is not mentioned in the SPS Agreement.²⁵⁶ As a result, the Appellate Body Report sheds little or no light on what should be the appropriate content of WTO findings in connection with risk management.

One example from the hormones dispute illustrates the interaction between risk management and risk assessment. The example is whether the sanitary measures of the European Communities were inconsistent with SPS article 5.5, which proscribes "arbitrary or unjustifiable distinctions in the levels [a Member] considers to be appropriate in different situations."²⁵⁷ As discussed above,²⁵⁸ this Agreement provision appears to require a complicated analysis. However, the Panel and Appellate Body reports do not clearly indicate which factors are relevant and may be taken into account, and what standard or priority should be used to balance those factors.

It is not clear from the Panel's report what factors it considered relevant to deciding whether a difference in levels of protection is "arbitrary or unjustifiable." For example, the Panel does not clarify the extent to which different measures could be justified by risk-benefit analyses that take into account anticipated benefits²⁵⁹ or consumer perceptions of risk.²⁶⁰ But if

253. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.100, 8.163.

254. See, e.g., *id.* ¶ 8.97 ("an assessment of risks is . . . a scientific examination of data and factual studies; it is not a policy exercise . . ."); *id.* ¶ 8.163 ("[T]here is a distinction between risk assessment which is a scientific examination and risk management which involves social value judgments").

255. The Hormones Panel employed a bright-line distinction between risk assessment and risk management:

there is a distinction between *risk assessment* which is a scientific examination and *risk management* which involves social value judgments. Once the risks have been assessed, i.e., once the risks and their probability of occurrence identified, a Member will need to decide, on the basis of its own value judgments, whether it can accept these risks. In so doing a Member sets its "appropriate level of sanitary protection". The determination and application of the appropriate level of protection by a Member is part of risk management.

Hormones Panel Report (CAN), *supra* note 13, ¶ 8.163.

256. See *supra* text accompanying note 244.

257. The Panel found that the Committee on Sanitary and Phytosanitary Measures had yet to develop guidelines to help implement this provision. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.173.

258. See *supra* Part I.B.1.

259. The relevance of benefits remains unclear. On the one hand, the Panel at times seemed to consider that the only justification for a difference in levels of protection must be safety-based. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.216-8.217. On the other hand, the Panel apparently took into account "that, according to scientific experts advising the Panel, the hormones at issue . . . may also have beneficial effects

benefits could not be weighed in the balance, or consumer anxieties could not be respected, or domestic politics could not be taken into account, what would remain of the sovereignty inherent in risk management decisions? What are the permissible factors and what are the permissible balances that *would* justify differences in levels of protection? The Appellate Body Report is only slightly clearer than the Panel reports. One *might* infer from that Report that the nature of the regulatory intervention required might be used to justify different approaches,²⁶¹ or that the administrative difficulty and the cost of control might constitute a justification,²⁶² or that information about benefits as well as about risks might be relevant to a justification.²⁶³ Such inferences from the Appellate Body's Report would be indirect, however, and would undoubtedly be subject to revision in future cases.²⁶⁴ Therefore, we are left with little guidance on relevant fac-

(such as improved composition of the carcass upon treatment in terms of more lean meat and less fat)." *Id.* ¶ 8.235.

260. Canada argued against the measures of the European Communities in part because "the EC ban causing the distinction in levels of protection is based on additional factors not relevant to the protection of health (such as . . . meeting consumer anxieties and expectations . . .)." *Id.* ¶ 8.242. The Panel apparently thought that "consumer preferences" are irrelevant in risk assessment. *Id.* ¶ 8.108. The SPS Agreement itself, however, is clear in regarding the voluntariness of exposure as a factor relevant to justifying a difference in levels of protection. SPS Agreement, *supra* note 3, art. 5.5. The justifying role of consumer perceptions or anxieties remains remarkably unclear in the Panel Report. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.181, 8.192, 8.216, 8.232-8.241.

261. The Appellate Body disagreed with the Panel's conclusion that the difference in the EC's levels of protection between "added hormones in treated meat" ("no-residue") and "naturally-occurring hormones in food" ("unlimited-residue") was arbitrary and unjustifiable. Appellate Body Report, *supra* note 16, ¶ 221. The Appellate Body found a "fundamental distinction" between added and naturally-occurring hormones, and stated with respect to naturally-occurring hormones that:

the European Communities simply takes no regulatory action; to require it to prohibit totally the production and consumption of such foods or to limit the residues of naturally-occurring hormones in food, entails such a comprehensive and massive governmental intervention in nature and in the ordinary lives of people as to reduce the comparison itself to an absurdity.

Id.

262. See *id.* ¶¶ 222-25 (concluding that the difference in the EC's levels of protection between hormones used for growth promotion and hormones used for therapeutic and zootechnical purposes was not "in itself" arbitrary or unjustifiable).

263. See *id.* ¶¶ 226-35 (agreeing with the Panel's conclusion that the difference in the EC's levels of protection between hormones used for growth promotion and the anti-microbial agents carbadox and olaquinox was unjustifiable under SPS article 5.5, but reaching this conclusion after considering arguments about therapeutic benefits and equally useful available alternatives to carbadox or olaquinox).

264. In a different but related context, the Appellate Body cited consumer anxieties in Europe in reversing the Panel's finding that the unjustifiable difference in the EC levels of protection between hormones used for growth purposes and the anti-microbial agents carbadox and olaquinox resulted in a hormone measure that constitutes "discrimination or a disguised restriction on international trade" inconsistent with SPS article 5.5. *Id.* ¶¶ 242-46; SPS Agreement, *supra* note 3, art. 5.5. The Appellate Body Report recites that the documentation surrounding the EC's adoption of its hormones measures:

makes clear the depth and extent of the anxieties experienced by the European Communities concerning the results of the general scientific studies (showing the carcinogenicity of hormones), the dangers of abuse (highlighted by scandals

tors and no standard by which to balance competing factors.

Fortunately, it is not the objective of this Article to sort out a theory of justifiable inconsistency, but to focus on the roles of scientific uncertainty and science policy in risk management, and on the nature of WTO factfinding concerning them. The task of comparing the risks posed by two different compounds illustrates the problems. Canada argued, for example, that the risks from the use of carbadox (an anti-microbial agent allowed by the EC as a feed additive in swine production) were at least as serious as those posed by the hormones in dispute, because carbadox is carcinogenic and mutagenic.²⁶⁵ The EC contended, among other things, that there was no risk under either regulatory regime because carbadox is used only in small quantities and is hardly absorbed, so that it leaves no residues in meat destined for human consumption.²⁶⁶ The Panel relied upon the advice of its experts in making factual findings against the European Communities on this question of the extent of the risk.²⁶⁷ Although

relating to black-marketing and smuggling of prohibited veterinary drugs in the European Communities) of hormones and other substances used for growth promotion and the intense concern of consumers within the European Communities over the quality and the drug-free character of the meat available in its internal market.

Appellate Body Report, *supra* note 16, ¶ 245. The Appellate Body rejected the Panel's apparent conclusion that the EC hormone measures "were not really designed to protect its population from the risk of cancer." *Id.* The Appellate Body held that the Panel's finding of "discrimination or a disguised restriction on international trade" was erroneous as a matter of law. *Id.*

Two things are noteworthy about the Appellate Body's reasoning. The first is that it was based on consumer anxiety about the risk of cancer, not on the risk of cancer itself. The Appellate Body had already upheld the Panel's finding that there was no adequate risk assessment to warrant the EC's hormone measures. *Id.* ¶ 208. Surely a documented risk of cancer would have been sufficient. But under SPS article 5.5, in the context of a member's objective or purpose in adopting a measure, consumer anxiety and actual risk are *both* relevant.

The second noteworthy feature is that consumer anxiety is apparently highly relevant in determining whether a measure is a "disguised restriction on international trade," but not in determining whether a difference in levels of protection is "arbitrary or unjustifiable." With respect to the difference in the EC levels of protection between growth promotion hormone use and use of carbadox and olaquinox, the Appellate Body upheld the Panel on the latter issue, prior to reversing it on the former. *Compare id.* ¶ 235, with *id.* ¶ 246. Yet these documented consumer anxieties seemed to play no role in the Appellate Body's reasoning on whether the difference in level of protection was justifiable. The regime now in place as a result of this case appears to be that even if an adequate risk assessment is not available, unjustifiable differences in levels of protection may be consistent with article 5.5 of the Agreement if the member has acted out of concern for consumer anxiety. *Id.* ¶¶ 208, 246. Such a curious result will likely be addressed in future cases.

265. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.228, 8.231-8.232, 8.236, 8.238-8.239, 4.226-4.275; Appellate Body Report, *supra* note 16, ¶ 226.

266. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.228, 8.231-8.232, 8.236, 8.238-8.239, 4.226-4.275; Appellate Body Report, *supra* note 16, ¶¶ 230, 232-233.

267. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶ 8.236 ("The experts also stated that additives in feedstuffs pose additional risks in that they may harm the persons handling the feedstuff."); *id.* ¶ 8.238 ("[A]ccording to the experts advising the Panel, there is no guarantee that the piglets treated with carbadox will not be slaughtered."); *id.* ¶ 8.239 ("[A]ccording to the experts advising the Panel, once a substance has

the Panel wished to claim that it was not itself engaging in the assessment of risks, it is not plausible that a panel can avoid doing so in such a case. Comparisons between risks posed by different compounds or by different regulatory measures are bound to be raised for the first time before WTO panels, without prior scientific meetings to address those precise issues. It is unreasonable to suggest that WTO panels can make comparisons without actually assessing the risks of the compounds or measures being compared. In the hormones dispute, the Appellate Body Report clearly condoned the Panel's factfinding concerning risk,²⁶⁸ and it is hard to see how the Appellate Body could not do so.

Moreover, future comparisons are likely to be far more complicated and far more uncertain than in the hormones dispute.²⁶⁹ For example, WTO panels may be called upon to decide whether carcinogenicity data showing only liver tumor effects in laboratory mice should be considered as evidencing a risk comparable to that posed by a compound with data showing tumor effects in multiple organs and two species.²⁷⁰ In the hormones dispute, the Panel probably believed that it would be relatively simple to make findings of fact about exposure and comparative risk in that context.²⁷¹ However, the Panel misperceived the implications of its undertaking. Although panels should not themselves conduct formal risk assessments, they are called upon to determine whether the science-policy choices of a member are scientifically plausible, in light of the scientific uncertainty present in the factual issues. Such findings are appropriate and necessary not only within the confines of risk assessment, but also with regard to the decision-making process that is essential to risk management.

been administered to an animal there will always be some residue level of this substance or a metabolite left, albeit a very small one, in the meat of that animal.").

268. Appellate Body Report, *supra* note 16, ¶¶ 220-35. For example, the Appellate Body noted and apparently relied upon the fact that the experts advising the Panel confirmed that carbadox is genotoxic, *id.* ¶ 226, the Panel's conclusion that the use of carbadox "poses additional risks since it may harm the persons handling the feedstuff," *id.* ¶ 230, and the Panel's statement "that, according to the experts advising it, once a substance has been administered to an animal, there will always be some residue of this substance or a metabolite . . . in the meat of that animal," *id.* ¶ 233.

269. Cf. Wirth & Silbergeld, *supra* note 88, at 1865-79 (discussing conceptual difficulties with comparative risk assessment, especially in context of scientific uncertainty).

270. This example has been chosen as illustrative of the role of science policy. See *supra* Part I.A.2.

271. The Hormones Panel also decided to rely upon the factual advice of its experts at other places within the Panel's review of the risk management decisions of the European Communities. See, e.g., Hormones Panel Report (CAN), *supra* note 13, ¶ 8.190 ("[A]ll scientific experts advising the Panel have concluded that residues of the three natural hormones present endogenously in meat and other foods or administered for therapeutic or zootechnical purposes are qualitatively the same as the residues of these hormones administered for growth promotion . . ."); *id.* ¶ 8.202 ("[A]ccording to scientific experts advising the Panel, zootechnical use of these hormones can occur on a large scale and at regular intervals, namely each year for oestrus synchronization of entire herds.").

B. The Evidence Properly Available to the Hormones Panel

The Hormones Panel exhibited an ambivalence about what evidence it should consider. On the one hand, the Panel allowed the parties to submit written scientific evidence.²⁷² But insofar as that evidence was “new evidence” produced before the Panel but not taken into account in any already completed risk assessment document, the Panel did not consider it as meeting the Panel’s procedural and substantive requirements for a risk assessment under SPS article 5.1.²⁷³ The Panel considered this position to be consistent with its view that it had no mandate to conduct its own risk assessment.²⁷⁴ Moreover, when the European Communities requested that the Panel consider the studies and other relevant data that Canada and the United States had used to approve the use of the hormones, the Panel “did not consider this information to be relevant to address the EC measures in dispute.”²⁷⁵ Nevertheless, as noted above, the Panel at various times based its findings of fact in part upon the opinions and information provided by its own experts.²⁷⁶

The Appellate Body’s position on this issue is not entirely clear, although its position is different than that of the Panel. The Appellate Body rejected the Panel’s “minimum procedural requirement” under article 5.1,²⁷⁷ and in so doing expressed concern that such a requirement “could well lead to the elimination or disregard of available scientific evidence that rationally supports the SPS measure being examined.”²⁷⁸ Moreover, in reviewing the evidence in the record to determine whether there was reasonable scientific support for the EC measures, the Appellate Body Report did not distinguish between older evidence and “new evidence” produced in the context of the Panel proceeding.²⁷⁹ One can infer, therefore, that “new evidence” produced for the first time in a panel proceeding can be used to find a plausible scientific basis for a member’s SPS measures.

Finally, the Panel was attempting an impossible task – making determinations about risk assessment documents without actually understanding or taking a position on risk assessment issues.²⁸⁰ There are sound policy reasons why WTO panels should make independent findings about the merits of risk assessments. The Hormones Panel should have used all of the relevant scientific information made available to it by the parties and by its own appointed experts, but should have used it for the limited purpose of finding whether there was any reasonable scientific basis for the risk determinations of the European Communities.

272. *Id.* ¶ 8.10.

273. *Id.* ¶ 8.118.

274. *Id.*

275. *Id.* ¶ 8.11; Hormones Panel Report (US), *supra* note 13, ¶ 8.11.

276. *See supra* notes 231, 267, 271.

277. Appellate Body Report, *supra* note 16, ¶¶ 188-91.

278. *Id.* ¶ 190.

279. *See id.* ¶¶ 192-209; Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.111, 8.118, 8.133-8.136.

280. *See supra* Part III.A.1.

C. The Standard and Burden of Proof Employed by the Panel

The Panel surely employed some standard of proof in reaching its findings of fact, but that standard is not explained in its report. The Panel made findings about the weight of evidence produced by the parties.²⁸¹ Moreover, the Panel in a number of instances relied upon the advice of its experts and made findings of fact.²⁸² But the Panel did not state the standard of proof it employed.²⁸³ The Panel did not address the question, "How probative or convincing does the evidence have to be in order to warrant a finding in favor of a proponent on any factual issue?"

The Appellate Body Report also provided no direct answer to the question. The Appellate Body discussed the appropriate "standard of review" for a WTO panel to use in reviewing a member's acts, determinations, or evidence.²⁸⁴ It pointed out that neither the SPS Agreement, the DSU, nor any covered agreements (other than the Anti-Dumping Agreement) prescribes a particular standard of review.²⁸⁵ Therefore, neither a panel nor the Appellate Body is authorized to adopt a standard of review that might change the "finely drawn balance" of the Agreement.²⁸⁶ However, Article 11 of the DSU requires a panel to make "an objective assessment of the matter before it, including an objective assessment of the facts of the case and the applicability of and conformity with the relevant covered agreements. . . ." ²⁸⁷ This "duty to make an objective assessment of the facts" was interpreted by the Appellate Body as a duty of "objectivity" in the

281. E.g., *Hormones Panel Report (CAN)*, *supra* note 13, ¶ 8.149 ("[W]e note that the European Communities has not provided convincing evidence that the control . . . of the hormones in dispute is more difficult than the control of other veterinary drugs the use of which it allows.").

282. See *supra* Parts III.A, III.B. An extraordinary example is provided by the finding in the *Hormones Panel Report*:

Addressing the fourth EC argument that there are no alternatives to carbadox or olaquinox available which have the same therapeutic action, we note that *one of the experts* advising the Panel stated that there are readily available alternatives, such as oxytetracycline. We thus find that this EC argument does not justify the distinction in levels of protection for the five hormones at issue when used as growth promoters and carbadox or olaquinox.

Hormones Panel Report (CAN), *supra* note 13, ¶ 8.237 (emphasis added). Whether this finding is viewed as being about the comparative therapeutic effects of alternative treatments or about the economic availability of market substitutes or about the weight of a cost-benefit justification for alternative regulatory measures, it is remarkable for its slim evidentiary basis and for what it shows about the nature of the subtle but important factfinding implicit in these WTO panel reports.

283. I argued above that a panel should make its findings by a preponderance of all the evidence. See *supra* Part II.C. At least the Panel seemed to be of the view that "science can never provide a certainty, i.e. exclude once and for all that a specific substance can ever have adverse health effects." See *Hormones Panel Report (CAN)*, *supra* note 13, ¶ 8.155.

284. Appellate Body Report, *supra* note 16, ¶¶ 110-19.

285. *Id.* ¶ 114; see also Steven P. Croley & John H. Jackson, *WTO Dispute Procedures, Standard of Review, and Deference to National Governments*, 90 AM. J. INT'L L. 193, 199 (1996) (noting that no provisions in the DSU explicitly concern the "standard of review" as such).

286. Appellate Body Report, *supra* note 16, ¶ 115. See *id.* ¶ 117.

287. *Id.* ¶ 116; DSU, *supra* note 7, art. 11.

sense of "good faith."²⁸⁸ This duty requires a panel to consider the evidence presented, to make its factual findings on the basis of that evidence, and to afford the party submitting the evidence "fundamental fairness, or what in many jurisdictions is known as due process of law or natural justice."²⁸⁹ The Appellate Body cannot reverse a factual determination of a panel merely because there has been an error of judgment in the appreciation of evidence,²⁹⁰ but only if there has been "an egregious error that calls into question the good faith of a panel."²⁹¹

This discussion by the Appellate Body, however, does not answer the question presented here. How probative or convincing does the proponent's evidence have to be in order to warrant or compel a finding of fact in the proponent's favor? It is no answer to say that a panel should consider the evidence "objectively" and in good faith. Nor is it responsive to add that it is generally within a panel's discretion both to decide which evidence to rely upon in making findings²⁹² and to determine the credibility and weight properly to be ascribed to a given piece of evidence.²⁹³ The proper standard of proof should be a matter of law binding upon all panels, not decided by individual panels on a case-by-case or issue-by-issue basis. I have argued elsewhere that the preponderance standard of proof is the standard that creates an incentive for parties on both sides of an issue to produce adequate evidence, and that this standard treats all parties in an unbiased fashion.²⁹⁴ It is possible to interpret "an objective assessment of the facts" as encompassing both of these policy objectives, and as imposing upon panels an obligation to make findings in accordance with the preponderance of the evidence. Neither the Hormones Panel nor the Appellate Body has chosen this interpretation, however, and so it appears that future panels are free to find facts by a preponderance of the evidence or by some higher or lower standard.

There is some indication, however, that the Appellate Body is open to developing an enforceable notion of the minimal evidentiary support needed for a panel's findings. It should be inconsistent with "an objective assessment of the facts" to make a finding when the panel record does not contain that minimum of evidence which any reasonable person would consider necessary to support such a finding.²⁹⁵ For example, the Appellate Body reversed the Panel's finding that the EC's different levels of protection for the growth-promotion use of hormones and the anti-microbial

288. Appellate Body Report, *supra* note 16, ¶ 133.

289. *Id.*

290. The DSU limits appellate review to "issues of law covered in the panel report and legal interpretations developed by the panel." DSU, *supra* note 7, art. 17.6. See also Appellate Body Report, *supra* note 16, ¶ 132 ("Findings of fact, as distinguished from legal interpretations or legal conclusions, by a panel are, in principle, not subject to review by the Appellate Body.").

291. Appellate Body Report, *supra* note 16, ¶ 133.

292. *Id.* ¶ 135.

293. *Id.* ¶ 132.

294. Walker, *supra* note 169, at 1113-20.

295. For the discussion of the requirement of minimal rational support, see *supra* Part II.C.

use of carbadox and olaquinox resulted in discrimination or a disguised restriction on international trade.²⁹⁶ Part of the basis for this reversal was that, even if a difference in levels of protection is unjustifiable, the mere degree of difference between a "no residue" level and an "unlimited residue" level is not sufficient evidence to support such a finding.²⁹⁷ The remainder of the evidence was likewise insufficient to support the Panel's finding.²⁹⁸ Thus, the Appellate Body held that the Panel's finding was erroneous as a matter of law.²⁹⁹ Under its authority to interpret the DSU ("objective assessment") and to interpret specific provisions of covered agreements, the Appellate Body has the authority and should have the responsibility to determine whether a panel's evidence provided the necessary rational support for its findings.

The allocation of the burdens of production and of persuasion are extremely important in a case like the growth hormones dispute, where there may be significant scientific uncertainty about carcinogenic risk. Especially where the factual issue is posed as a yes-or-no question of whether the relevant residues are "safe" or "unsafe," the party with the burden of persuasion has a difficult burden of proof, even under a preponderance standard.³⁰⁰ The reason is that in all likelihood no one knows for certain whether the residues are safe or unsafe, and there is some difference of scientific opinion on whether the residues are probably safe. However, with certain assumptions and models risk assessors can make plausible scientific estimates of the risk and can explain the uncertainties inherent in those estimates. A factfinding panel is understandably tempted to resolve its most difficult factual issues by assigning the burdens of production and persuasion.³⁰¹

The Panel correctly found that the initial burden rests on the complaining party to present a *prima facie* case of an inconsistency between a sanitary measure and the SPS Agreement.³⁰² It also correctly interpreted a "*prima facie* case" as being "factual and legal arguments that, if unrebutted, would demonstrate a violation of the SPS Agreement."³⁰³ The Appellate Body took a slightly different approach in defining a "*prima facie* case" as

296. Appellate Body Report, *supra* note 16, ¶¶ 236-46.

297. *Id.* ¶ 240.

298. *Id.* ¶ 246.

299. *Id.*

300. A qualitative, yes-or-no question of "safety" is probably not the best way to frame the issue, however. What is needed is a more tractable series of issues, with a clear allocation of burdens of production and persuasion under each. Then a panel could determine the extent of scientific uncertainty under each issue and sub-issue, as well as the scientific plausibility of the science-policy choices made by members.

301. For an illustration, see the Hormones Panel Report (CAN), *supra* note 13, ¶ 8.237, quoted in its entirety *supra* note 282. One reason that seemed to warrant making this finding on such a slim evidentiary basis was that the Panel had placed the burden of providing a justification on the European Communities, *id.* ¶ 8.241, and had found that the European Communities had "not submitted scientific evidence in support of these alleged justifications," *id.* ¶ 8.233.

302. *Id.* ¶ 8.54; Appellate Body Report, *supra* note 16, ¶ 98.

303. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.54.

“one which, in the absence of effective refutation by the defending party, requires a panel, as a matter of law, to rule in favor of the complaining party presenting the *prima facie* case.”³⁰⁴ The difference between the two notions is subtle but important, and involves the role of the Appellate Body.

Under the Panel’s definition, a panel might find as a matter of fact that it considers the evidence presented to be sufficiently weighty, so that in the absence of contrary evidence it would find for the complaining party. On this view, calling the presented evidence a “*prima facie* case” merely means that the factfinder considers it convincing, at least as an initial matter and ignoring any rebuttal evidence. Under the Appellate Body’s definition, however, a complaining party’s evidence does not constitute a *prima facie* case unless it is so compelling that the factfinder has no authority to rule against it in the absence of contrary evidence. On this latter view, an un rebutted *prima facie* case entitles the Appellate Body to reverse a panel that finds against the complaining party. Presumably, evidence that is merely persuasive to the panel, but which is not so compelling as to authorize Appellate Body reversal, is not a *prima facie* case. A true *prima facie* case creates a rebuttable presumption for the complaining party that is enforceable by the Appellate Body.

This difference in definition may be subtle, but the procedural implications can be extremely important. Under the Panel’s definition, calling the evidence presented a *prima facie* case is nothing more than an announcement that the panel finds that evidence persuasive if it remains un rebutted. Doing so informs the defending party that the plaintiff’s evidence appears strong, and that the defendant would be well advised to produce contrary evidence. Under the Appellate Body’s definition, however, it can be said that a *prima facie* case shifts a true burden of *production* to the defending party. Absent rebuttal evidence, the plaintiff should prevail as a matter of law. I have argued above³⁰⁵ for a burden of production and a definition of *prima facie* case intermediate to those of the Hormones Panel and of the Appellate Body. A complaining party should have to produce the minimally sufficient evidence needed to sustain a finding in its favor, and such sufficiency would be subject to Appellate Body review. The complaining party, however, should not have to produce evidence so strong as to compel such a finding as a matter of law. Under my approach, a *prima facie* case would in fact be persuasive to the panel, but would not in itself shift a burden of production to the defendant. The defendant would therefore be entitled to rely merely on argument to dissuade the panel from finding for the complaining party.

The choice among the three definitions may influence whether the defending member has a burden of production, but the choice should make no difference in the allocation of the burden of *persuasion*. The Panel incorrectly found that once such a *prima facie* case has been presented, the

304. Appellate Body Report, *supra* note 16, ¶ 104.

305. See *supra* text accompanying notes 184-91.

burden of persuasion shifts to the defending party.³⁰⁶ The Appellate Body Report is more ambiguous about which burden shifts.³⁰⁷ I have argued above,³⁰⁸ however, that if both parties produce minimally sufficient evidence for their allegations, a panel should weigh all the evidence and make findings in accordance with the preponderance of that evidence. In weighing all the evidence, a panel should still follow the default rule that the complaining party has the burden of persuasion. If a panel finds equally weighty evidence for and against an inconsistency with the Agreement, it should enter a finding against the complaining party. This is because an inconsistency has not been proved simply by the presentation of a *prima facie* case; if rebuttal evidence is produced, an inconsistency has not been proved unless the complaining party's evidence is more probative and convincing than the defending party's evidence.

Several examples from the hormones dispute illustrate the importance of how the burden of persuasion is allocated. The first involves the harmonization provisions of article 3. The Panel concluded that once a complaining party provides a *prima facie* case to demonstrate that a relevant international standard exists and that the measure in dispute does not conform to it,³⁰⁹ then the defending party must prove that its measure meets

306. As the Panel stated: "... Canada bears the burden of presenting a *prima facie* case of inconsistency with the SPS Agreement, after which the burden of proof shifts to the European Communities to demonstrate that its measures in dispute meet the requirements imposed by the SPS Agreement." Hormones Panel Report (CAN), *supra* note 13, ¶ 8.58.

It does make sense to say, however, that as a tactical matter, once a *prima facie* case has been produced by the complaining party, the defending party would be well advised to produce rebuttal evidence, because otherwise the unrebutted *prima facie* evidence will probably carry the day. See *supra* Part II.C; Measure Affecting Imports of Woven Wool Shirts, *supra* note 126, at 16-17 (once India had put forward a *prima facie* case to demonstrate that a U.S. measure was inconsistent with the Agreement on Textiles and Clothing, the United States risked an adverse finding by default if it failed to bring forward evidence sufficient to disprove India's claim).

307. The Appellate Body Report stated: "When that *prima facie* case is made, the burden of proof moves to the defending party, which must in turn counter or refute the claimed inconsistency [with the SPS Agreement]." Appellate Body Report, *supra* note 16, ¶ 98. The problem comes in specifying the meaning of "counter" and "refute." If these words merely mean "produce the minimum weight of evidence needed to warrant a finding in its favor," then all that shifts is a burden of production. If, however, these words mean "produce evidence whose probative value outweighs that produced by the complaining party," then the defending party would have the burden of persuading the panel that no violation in fact occurred. The Appellate Body Report also noted that: "Only after such a *prima facie* determination had been made by the Panel may the onus be shifted to the European Communities to bring forward evidence and arguments to prove the complaining parties claim." *Id.* ¶ 109. This statement contains the same ambiguities in the words "bring forward . . . to prove . . ." There would seem to be no justification at all for shifting to the defendant an even weightier burden than proving by a preponderance — namely, the burden of producing evidence so compelling that the panel must find in the defendant's favor as a matter of law.

308. See *supra* Part II.C.

309. The Panel identified the "conform to" language of SPS article 3.2 with the "based on" language of SPS article 3.1, and afforded the same presumptions created in favor of conforming measures also to measures that are *based on* international standards. See SPS Agreement, *supra* note 3, arts. 3.1, 3.2; Hormones Panel Report (CAN), *supra* note 13, ¶

the provisions of SPS article 3.3.³¹⁰ The Panel also interpreted both conditions under SPS article 3.3 as requiring compliance with the risk assessment and risk management provisions of SPS article 5.³¹¹ As a result, the Panel found that, in such a case, the defending party bears the burdens of production and persuasion in proving that the measure complies with the relevant article 5 provisions.³¹² Once the Panel found, therefore, that the measures of the European Communities did not conform to existing international standards,³¹³ it shifted the burden of persuasion to the European Communities to prove compliance with SPS article 5.³¹⁴

8.89 ("Article 3.2 . . . specifies that the complaining party has the burden of overcoming a presumption of consistency with the SPS Agreement in the case of a measure *based on* international standards.") (emphasis added); Appellate Body Report, *supra* note 16, ¶ 162. Such a collapse of SPS articles 3.2 and 3.1 is inappropriate. See *supra* Part I.C.

310. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.90. The Panel interpreted SPS article 3.1 as imposing a general obligation on all members to base all sanitary measures on international standards, guidelines and recommendations (when such standards, guidelines and recommendations exist), and SPS article 3.3 as providing an exception to that general obligation. *Id.* ¶ 8.89. Reading SPS article 3.3 as an exception to article 3.1 is inappropriate. Appellate Body Report, *supra* note 16, ¶ 172.

311. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.86. I argued above that this collapse of the two alternative conditions of SPS article 3.3 is unnecessary. See *supra* Part I.C.

312. Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.92-8.93.

313. This part of the Panel's reasoning is also puzzling. The Panel found that the European Communities' measures represented a level of protection "significantly *different* from the level of protection set by the Codex standards" and were therefore not "based on" existing international standards. *Id.* ¶¶ 8.78-8.80. The Panel then made the peculiar statement that "[f]or purposes of our analysis under Article 3.3, we assume that the former level [the level of protection of the European Communities] is *higher* than the latter [the Codex level]." *Id.* ¶ 8.83. Why should the Panel "assume" this to be true rather than find it to be true?

Perhaps this anomaly is explained by the fact that the complaining members were in effect claiming that the level of protection of the European Communities was *not* higher than that provided by the Codex recommendations. If the dispute is conceptualized in terms of "safety," then the complaining parties were in effect claiming that the Codex standards were "safe" and that the measures of the European Communities were no safer. But in that case, why would the measures not be "based on" or not "conform to" the international guidelines? The Panel elsewhere seemed to agree with its expert advisors that:

the fact that ADIs and MRLs exist for zeranol and trenbolone and not for the natural hormones does not . . . *per se* mean that the latter are inherently safer than the former since the international standards for both synthetic and natural hormones reflect essentially the same level of protection, namely a "no appreciable risk" level.

Id. ¶ 8.216.

The reasonable solution may be that SPS article 3.3 should not have applied at all, and that the dispute should have been resolved under SPS article 2.2 (a measure must be applied "only to the extent necessary to protect human . . . health") and SPS article 5.6 (measures should be "not more trade-restrictive than required to achieve their appropriate level of sanitary . . . protection"). SPS Agreement, *supra* note 3, arts. 2.2, 3.3, 5.6. But if that is so, then the entire analysis of the Panel was largely misdirected. See Hormones Panel Report (CAN), *supra* note 13, ¶ 8.274 (further findings under SPS art. 2 unnecessary); *id.* ¶ 8.250 (findings under SPS art. 5.6 unnecessary). Additionally, the burden of proof analysis of SPS articles 3.1 and 3.3 was gratuitous.

314. Remarkably, even in the case of the hormone MGA, for which there was no international standard, the Panel shifted the burden of persuasion to the European Commu-

The Appellate Body explicitly found that the Panel erred when it collapsed the concepts of "conform to" (article 3.2) and "based on" (article 3.1), and held that the SPS Agreement did not impose on members a basic obligation to conform domestic sanitary measures to Codex standards, guidelines and recommendations.³¹⁵ However, if proof of a lack of conformity is no longer a *prima facie* case, then what showing by a complaining party is sufficient to create a rebuttable presumption that the challenged measures are inconsistent with the Agreement? In the case of the five hormones where Codex standards existed, the Appellate Body assumed that the EC measures were not "based on" Codex standards and that they resulted in a higher level of protection than the Codex standards would provide.³¹⁶ Therefore, the Appellate Body agreed with the Panel that the challenged measures fell under article 3.3 and that the European Communities had an obligation to comply with the risk assessment requirement of article 5.1.³¹⁷

The initial burdens of production and persuasion concerning a violation of article 5.1 remained on the United States and Canada.³¹⁸ The Appellate Body found, however, that these parties had in fact produced a *prima facie* case that the EC measures were *not* based on a risk assessment.³¹⁹ The Appellate Body reached this conclusion after reviewing the proffered evidence on its merits.³²⁰ Although it is not at all clear what generalizable rule might be inferred from this treatment, the Appellate Body seemed to be persuaded by the Panel's finding that "[a]ll of the scientific studies" that specifically addressed the issue concluded that the use of the hormones for growth purposes is "safe," as well as by the lack of specificity in the contrary scientific evidence.³²¹ Presumably, given the Appellate Body's definition of "*prima facie* case," the evidence produced by the United States and Canada was so compelling that it entitled them to a finding that the measures were inconsistent with the Agreement as a matter of law, in the absence of "effective refutation" by the EC.³²² Throughout this battle, the evidence remained the same, but the burden of persuasion played an important role in the outcome.

In a second example bearing on risk assessment, the shifted burden of persuasion affected factfinding under the Panel's substantive requirements for risk assessment under article 5.1.³²³ The Panel found that the European Communities had "the burden of proving the existence of a risk

nities. See Hormones Panel Report (CAN), *supra* note 13, ¶ 8.256 (*prima facie* case of violation under SPS article 5.1-5.3, dealing with risk assessment); *id.* ¶ 8.266 (*prima facie* case of violation under SPS article 5.5, dealing with differences in levels of protection).

315. Appellate Body Report, *supra* note 16, ¶¶ 163-66.

316. *Id.* ¶ 174.

317. *Id.* ¶¶ 176-77.

318. *Id.* ¶ 197 n.180.

319. *Id.*

320. See *id.* ¶¶ 195-200.

321. See *id.* ¶¶ 196-200; see *supra* text accompanying notes 213-20.

322. See *supra* text accompanying notes 302-07.

323. See *supra* Part III.A.1.

assessment (and, derived therefrom, an identifiable risk)" on which the EC's measures were based.³²⁴ Under the Panel's interpretation, the EC had the burden of proving that the scientific conclusion about risk reflected in the EC's measures was in conformity with any of those reached in the studies referred to by the EC.³²⁵ Thus, the Panel's factfinding deliberations were reduced largely to a hunt for eligible scientific quotations about "safety,"³²⁶ and produced the inconsistencies discussed earlier.³²⁷ Although the Appellate Body rejected the Panel's interpretation of article 5.1 and its approach to burden shifting,³²⁸ the Appellate Body still found a *prima facie* case in the panel record³²⁹ and arguably shifted to the EC the burden of persuasion in "refuting" that *prima facie* case.³³⁰ It is possible therefore that the EC had the difficult task of persuading the WTO that its evidence was sufficiently specific and probative so as to overcome a presumption of no risk.³³¹ In view of the lack of standards for what constitutes "specifically focused" evidence, this was a substantial burden indeed.³³²

As a final example, in the area of risk management, the Panel found that the European Communities had the burden of proving that the selection and implementation of its appropriate level of protection was consistent with SPS article 5.5.³³³ This placed on the European Communities the tasks of persuading the Panel as to the comparative risk posed by different regulatory measures³³⁴ and of persuading the Panel that some risk/benefit analysis "justified" the choice among different regulatory measures.³³⁵ In this area, where the issues and relevant factors are entirely unclear,³³⁶ this is a substantial burden. The burden on the defending party, together with the vagueness of the task, helps to explain the finding of the Hormones Panel against the EC.³³⁷ Although the Appellate Body reversed this find-

324. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.154.

325. *Id.* ¶¶ 8.162, 8.120. I have already argued that the Panel misconceived the content of the findings it was supposed to make, and that the Panel should have made findings about whether there was any reasonable scientific basis for the risk assessment conclusions reached by the European Communities. See *supra* Part III.A. Here, however, I consider only the effect of the Panel's allocation of the burden of persuasion on the Panel's factfinding process.

326. See Hormones Panel Report (CAN), *supra* note 13, ¶¶ 8.122-8.140.

327. See *supra* Part III.A.

328. See *supra* text accompanying notes 234-36.

329. Appellate Body Report, *supra* note 16, ¶ 197 n.180.

330. Although the Appellate Body's position on this issue is ambiguous, see *supra* note 307 and accompanying text, the Appellate Body certainly placed a heavy burden of some kind on the EC. See *supra* notes 211-19 and accompanying text.

331. See Appellate Body Report, *supra* note 16, ¶ 104 ("[A] *prima facie* case is one which, in the absence of effective refutation by the defending party, requires a panel, as a matter of law, to rule in favour of the complaining party presenting the *prima facie* case."); *supra* text accompanying notes 211-19.

332. See *id.*

333. Hormones Panel Report (CAN), *supra* note 13, ¶ 8.168.

334. See, e.g., *id.* ¶ 8.196.

335. See, e.g., *id.* ¶ 8.200.

336. See *supra* Parts I.B.1, II.A.2, III.A.2.

337. See Hormones Panel Report, *supra* note 13, ¶ 8.247.

ing,³³⁸ it did so by reviewing the interpretations, arguments and evidence relevant to the issue, and the issue of burden of persuasion did not play a significant role.³³⁹ The Appellate Body has held, however, that the initial burden of proof under article 5.5 is on the complaining party.³⁴⁰ Nonetheless, given the vagueness of the relevant standards,³⁴¹ the burden of persuasion may play a more determining role in future dispute settlement proceedings.

Conclusion

The confusion of the Hormones Panel and the errors it made are understandable in light of the complexity of the factual evidence, the difficulty of the numerous legal issues of first impression, and the time constraints on its proceedings. Although the Appellate Body Report correctly resolved certain errors committed by the Panel, it left other errors uncorrected and failed to clarify many other issues. Nevertheless, what remains is a blueprint for structuring future cases under the SPS Agreement.

I have argued that the factfinding approach being established for WTO dispute settlement proceedings is fundamentally flawed in a number of respects. Cases under the SPS Agreement implicate the momentous clash between the interest in efficient international trade and the sovereign duty to protect health. The WTO's structuring of the factfinding roles of its dispute settlement bodies does not promise to be a stable institutional solution for resolving such disputes. What is needed is a dispute settlement process organized so that a global consensus on the criteria for scientific plausibility can evolve. Such a process would assist in the gradual replacement of scientific uncertainty with scientific knowledge, but it would not elevate one group of scientists over another or adopt particular science policies and impose them on individual members.³⁴² This Article proposes

338. Appellate Body Report, *supra* note 16, ¶ 246.

339. *See id.* ¶¶ 210-46.

340. *Id.* ¶¶ 108-09.

341. *See supra* Parts I.B.1, II.A.2, III.A.2.

342. In this Article, I have not developed the critical institutional perspective needed to argue how the WTO can best achieve harmonization of sanitary measures under the SPS Agreement. Nevertheless, the framework developed here sets the stage for such an argument. A somewhat parallel situation occurred in a U.S. regulatory context in the early 1980s, when the National Research Council was charged with studying the possibility that a centralized scientific body was needed to conduct risk assessment for all federal regulatory agencies engaged in risk-based regulation, so as to reduce the influence of policy-makers on risk assessment. *See* NRC (1994), *supra* note 24, at 33. The NRC Committee rejected this proposal for various reasons, but recommended instead: that a clear conceptual distinction should be maintained between risk assessment and risk management, that the two activities should not be physically separated, and that agencies should develop and use inference guidelines (default assumptions) that make the role of science policy in risk assessment both principled and transparent. *See id.* at 33-34. A similar suggestion is possible in the WTO context. If WTO panels continue to elucidate the distinct roles of risk assessment and risk management under the SPS Agreement, and make findings concerning whether a science-policy option employed by a member has a reasonable scientific basis, then the WTO dispute settlement process will help create transparency in risk-based decision-making by members and will help

an approach that would promote a gradual development of consensus through the proper structuring of the WTO's factfinding process. By adopting this proposed approach, the WTO may avoid becoming the "World Trans-science Organization," an organization that would surely pose a threat to sovereignty, and perhaps even to human health.

create an incentive for members to collaborate in areas of significant scientific uncertainty. WTO panels should not establish science policies, but should identify reasonable science policies that have been adopted by members and reward the adoption of such policies by deferring to them.