To defer or not to defer? Has the Appellate Body resolved the issue of an appropriate standard of review in SPS cases?

Tracey Epps, Lecturer, Faculty of Law, University of Otago, New Zealand

1. Introduction

One of the most vexed issues in World Trade Organization (WTO) dispute settlement is what standard of review dispute panels ought to apply when assessing the consistency of a country's domestic regulatory measures with trade rules. This is particularly problematic in the context of social regulations such as those enacted to protect health. In this context, Professor Trebilcock has called for a relatively deferential standard of review. He recognizes, however, that too wide a degree of deference could have the undesired result of allowing countries free rein to take protectionist measures that undermine the multilateral trading system.[1]

The question of what standard of review to apply has arisen most acutely in the context of the *Agreement on the Application of Sanitary and Phytosanitary Measures* (SPS Agreement). The SPS Agreement requires countries to ensure that sanitary and phytosanitary (SPS) measures not based on international standards are not maintained without scientific evidence (Article 3.3) and are based on a risk assessment (Article 5.1). These requirements have been the subject of a number of disputes, leading to a growing body of jurisprudence that has, *inter alia*, considered the standard of review for panels to apply when reviewing the conduct and conclusion of countries' risk assessments. This can be a controversial question because of the SPS Agreement's application to politically sensitive land-based sectors and to important questions of public safety. On one hand, there tends to be high levels of protectionism in these sectors. Too wide a degree of deference might allow countries to cover up protectionism by fudging complex scientific facts to make it look as though there is a legitimate case for a trade restrictive measure. Conversely, governments have a legitimate interest in protecting the health of humans,

---

[1] {Trebilcock, 2002 #11 @ 541}

animals and plants. This interest augurs for a lesser standard of deference to ensure that they are able to take such measures when necessary in the face of risk.

This paper takes up the question of how to strike a balance between a degree of deference that accords countries freedom to regulate, while ensuring that they do not have open rein to justify any trade restrictive measure on health-related grounds. Section two provides an overview of the SPS Agreement. Section three explores the tension which is at the core of the question of how much deference to afford, namely, between trade liberalization and domestic regulatory autonomy. The standard of review adopted by panels is key in balancing these interests. It is of both procedural and substantive importance as it directly implicates the allocation of power between the WTO and national authorities.[2] Section four examines the relevant WTO case law, including in particular the decisions in *EC – Measures Affecting Meat and Meat Products (Hormones)* (*EC – Hormones*) and *US – Continued Suspension*. Section five discusses whether through these cases the Appellate Body has resolved the problem of what constitutes an appropriate standard of review in SPS cases, and thus, how much deference to pay to countries' decisions. Section six concludes, noting that the position taken by the Appellate Body thus far, while very appealing on its face, nevertheless fails to resolve the problem. While the Appellate Body advocates a certain degree of deference, panels have insufficient guidance as to how to go about reviewing complicated scientific evidence. Without such guidance, it suggests that panels will inevitably drift towards taking an approach that comes inappropriately close to second-guessing the scientific conclusions reached by national authorities, thus intruding too far on domestic regulatory autonomy.

2. The SPS Agreement

The SPS Agreement sets out various procedural and substantive standards to which countries must conform when enacting SPS measures. Such measures are those designed to protection human, animal or plant life or health from various risks including those

---

[2] Matthias Oesch, "Standards of Review in WTO Dispute Resolution" (2003) Journal of International Economic Law 6:3 635 at 636.

arising from additives, contaminants, toxins or disease-causing organisms in foods, beverages or feedstuffs and diseases carried by animals, plants or products thereof.[3] The SPS Agreement seeks to further the use of harmonized SPS measures, on the basis of international standards, guidelines and recommendations developed by relevant international organizations.[4] Measures conforming to relevant international standards are presumed to be in compliance with the SPS Agreement's requirements.

If members wish to introduce or maintain measures more stringent than provided for by relevant international standards, they must ensure that they are supported by a *scientific justification* (Article 3.3). They must ensure that any such measure is "applied only to the extent necessary to protect human, animal or plant life or health, is based on scientific principles and is not maintained without sufficient scientific evidence" (Article 2.2). Countries thus have freedom to choose their own substantive standards, while imposing some constraints to prevent the use of standards for protectionist purposes.

The scientific justification requirement is subject to Article 5.1 which requires that members conduct a *risk assessment.* It states that members must ensure their SPS measures are "based on an assessment, as appropriate to the circumstances, of the risks to human, animal or plant life or health, taking into account the risk assessment techniques developed by the relevant international organizations." Article 5.2 requires the risk assessment to take into account various factors, including "available scientific evidence".

The SPS Agreement recognizes, subject to various conditions, the right of Members to set their appropriate level of protection once scientific evidence has established the presence of a risk (Article 5.3). Conditions include requirements to: ensure that any discrimination is not arbitrary or unjustifiable and that measures are not "applied in a manner which would constitute a disguised restriction on international trade" (Article 2.3); to take into account the objective of minimizing negative trade effects (Article 5.4);

---

[3] SPS Agreement, Annex A.
[4] For food safety: Codex Alimentarius Commission ("Codex"); for animal health: the International Office of Epizootics ("OIE"); and for plant health: organizations operating within the framework of the International Plant Protection Convention ("IPPC").

and to ensure that SPS measures are "not more trade-restrictive than required to achieve their appropriate level of sanitary or phytosanitary protection, taking into account technical and economic feasibility" (Article 5.6).

3. Balancing competing objectives

A fundamental theme of international economic law is the underlying tension between trade liberalization objectives and domestic regulatory autonomy. As noted above, this tension is particularly evident in the case of the SPS Agreement due to governments' interest in protecting the health of their citizens and the viability of their agricultural and horticultural sectors. Doing so may require trade restrictive measures when imported products present risks to the health of humans, animal, or plants. Trade liberalization and its associated goals will as a consequence suffer a setback to the extent of the restrictive measure. The WTO seeks, *inter alia*, to enhance global and domestic welfare by maximizing the realizable benefits from unilateral and multilateral trade liberalization.[5] These goals are reflected in the Agreement Establishing the WTO[6] which emphasizes various objectives in its Preamble, including raising standards of living, ensuring full employment and growing volume of real income and effective demand, as well as sustainable development. However, if countries are prevented from enacting and maintaining SPS measures, domestic regulatory goals may be negatively affected where human, animal, or plant life or health is compromised.

When negotiating the SPS Agreement, countries were cognizant of the fact that in some cases, trade restrictions are inevitable to allow proper protection of human, animal, and plant life and health. The SPS Agreement thus represents a carefully negotiated balance between the interests of members in supporting trade liberalization and in maintaining

---

[5] The benefits of trade liberalization follow from countries specializing in producing the goods in which they have a comparative advantage – which arises where the opportunity cost of producing a good (in terms of other goods) is lower in the home country than it is in other countries. James Gerber, *International Economics*, 2 ed. (Boston: Addison Wesley, 2002) at 41.

[6] *Marrakesh Agreement Establishing the World Trade Organization*, 15 April 1994, in *The Legal Texts: The Results of the Uruguay Round of Multilateral Trade Negotiations* (Cambridge: Cambridge University Press, 1999).

significant domestic regulatory autonomy.[7] It implicitly assumes the benefits of trade (by aiming to minimize the negative effects of SPS measures on trade) yet also recognizes that discrimination may be necessary where goods from a particular source pose a greater risk than those from other sources.

Each set of goals (trade liberalization and domestic regulation) is concerned with enhancing welfare (broadly defined). The question is thus one of determining when the benefits of trade liberalization outweigh those of domestic regulatory action, and *vice versa*. Where welfare gains from the domestic measure would outweigh the benefits of trade liberalization, trade may be restricted. On the other hand, where the benefits of trade liberalization outweigh domestic regulatory gains, countries may not be permitted to take trade restrictive measures. This much is reasonably uncontroversial; however, determining where the balance of interests lie is far more difficult.

The above characterization implies a continuum of situations based upon the welfare interests at stake. At one end, trade interests give way to domestic regulatory autonomy, while at the other, domestic autonomy gives way to trade liberalization objectives. The archetypal case at one end of the continuum is where trade a measure is enacted purely to secure the protection of domestic producers – the interests of a country in enacting such a measure is clearly outweighed by trade liberalization interests and the measure will be disallowed under WTO rules.[8] At the other end is a measure that is undeniably necessary to protect health (a ban on imports of food products infected with *E-coli* for example). Such a measure has clear welfare benefits and may be permitted despite its trade restrictive effect. These are relatively simple cases in respect of which the SPS Agreement avoids tension between trade liberalization and domestic regulatory autonomy. It does so by making it clear which goal is to take priority and in this regard it relies on science. It provides that trade restrictive SPS measures will be tolerated provided they

---

[7] Michael J. Trebilcock & Robert Howse, *The Regulation of International Trade*, 3rd ed. (London: Routledge, 2005) at 134.

[8] Article III prohibits discrimination by governments against foreign importers in favour of domestic producers of like products.

are *necessary* to protect health, with necessity required to be proven by reference to scientific evidence.

Matters become more complicated in the middle of the continuum, where it is not clear where the balance of interests objectively lies. This is the case where it is not clear that there is actually a risk to health (for example, does genetic modification present a health risk?); it is not clear that a risk is serious enough or likely enough to outweigh losses from trade restrictions (for example, where a food additive presents a minor risk of allergic reactions); or different values among countries make agreement impossible as to whether a given measure will contribute to domestic welfare (for example, the European public sees a risk in the use of growth hormones in cattle, while the American public does not). In these and other similar cases, tension between trade and health objectives persists and dispute panels and the Appellate Body will face difficulties – both conceptual and factual – trying to balance trade liberalization and domestic regulatory autonomy.

This article is about what happens in cases that lie in the murky ground in the middle of the continuum, when disagreement between countries as to the welfare gains of a particular domestic SPS measure lead to a complaint under the WTO's dispute settlement system, and a dispute panel is called upon to adjudicate as to whether or not the measure is *necessary*. These are the cases where the question of how much deference panels ought to show to countries' regulatory decisions is at its most stark.

The goals of the WTO's dispute settlement system are noted in Article 3 of the Dispute Settlement Understanding (DSU), namely, to settle conflicts efficiently and to provide predictability and security in the multilateral trading system. By achieving these goals, the system contributes to the WTO's task of liberalizing trade.[9] Further, in the SPS context, the dispute settlement system should not only be efficient in terms of resolving disputes promptly, but also in promoting welfare by furthering both the goal of trade

---

[9] Chad P. Bown, "On the Economic Success of GATT/WTO Dispute Settlement" (2004) 86:3 The Review of Economics and Statistics 811 at 812.

liberalization and that of protecting countries' regulatory policy space. The SPS Agreement seeks a balance between these goals, thus the logical inference is that the dispute settlement system will further welfare most efficiently if it can effectively balance trade liberalization and domestic regulatory autonomy in a manner that maximizes total welfare, both domestic and global. As will be discussed in section five, the appropriate standard of review should in turn be viewed as advancing the same ends.

Determining the location of this balance between trade liberalization and domestic regulatory autonomy is no easy matter. The SPS Agreement speaks about measures being *necessary* to protect animal, plant, or human health. It uses science as a proxy to determine when a measure is necessary – if the regulating country can show that it based their measure on scientific evidence, necessity will be found, and measures may be allowed that restrict trade. In other words, the welfare benefits of the SPS measure are said to outweigh the benefits of trade liberalization whenever scientific evidence shows that the measure is *necessary* to protect life or health. However, in some cases it will be difficult, if not impossible, to determine by any objective scientific standard whether or not a measure is *necessary*. As I have discussed elsewhere,[10] science is more often than not uncertain and produces conflicting evidence, thus making it a blunt tool to balance the objectives of trade liberalization and domestic regulatory autonomy.

4. Deference accorded in WTO cases

When these murky cases are disputed under the WTO's DSU, panels are faced with the question of how much deference to pay to the regulating country's decision to enact the measure in question. Ought they to step back and give the benefit of the doubt to the regulating country? Or ought they to peer more closely into the domestic decision-making process?

---

[10] Tracey Epps, *International Trade and Health Protection: A Critical Assessment of the WTO's SPS Agreement* (Cheltenham: Edward Elgar, 2008).

The SPS Agreement is silent on the matter of an appropriate standard of review.[11] In *EC – Hormones*, the EC suggested two main alternative approaches. The first is "*de novo* review" which would allow a panel complete freedom to come to a different view from the competent authority of the regulating country. A panel would have to "verify whether the determination by the national authority was correct both factually and procedurally".[12] The second is described as "deference". The EC argued that under this standard, a panel should not seek to redo the investigation conducted by the national authority but instead examine whether the "procedure" required by the relevant WTO rules had been followed.[13]

The *EC – Hormones* dispute involved an EC Directive prohibiting the use of various hormones for growth promotion purposes in beef cattle.[14] The US and Canada complained, *inter alia*, that the EC Directive violated various provisions of the SPS Agreement, including the Article 5.1 requirement that measures be based on a risk assessment. The US and Canada were successful at both the panel and Appellate Body levels. The Appellate Body held that the EC had not applied risk assessment techniques to the particular risks that it claimed were the basis of its SPS measure. For five of the six hormones at issue, the EC did produce risk assessments that found a risk of cancer due to hormone exposure, but these risk assessments did not examine the risks associated with the particular hormones when used for growth promotion purposes. For the sixth hormone, no valid risk assessment existed at all.

The Appellate Body considered that the appropriate standard of review under the SPS Agreement is neither *de novo* review, nor "total deference". Rather it stated that the

---

[11] *European Communities - Measures Affecting Meat and Meat Products* (1998), WTO Doc. WT/DS26/AB/R, WT/DS48/AB/R (Appellate Body Report), at para 114. (*EC – Beef Hormones, AB Report*)
[12] EC's appellant submissions, para 122.
[13] *Ibid*. at para 123.
[14] *EC - Measures Concerning Meat and Meat Products (Hormones) Complaint by the United States* (1997), WTO Doc. WT/DS26/R/USA (Panel Report).*EC – Beef Hormones, AB Report*; *United States - Continued Suspension of Obligations in the EC-Hormones Dispute* (2008), WTO Doc. WT/DS320/R (Panel Report). (*US – Continued Suspension, Panel Report*); *United States – Continued Suspension of Obligations in the EC – Hormones Dispute* (2008), WTO Doc. WT/DS320/AB/R (Appellate Body Report). (*US – Continued Suspension, AB Report*)

standard should be "the objective assessment of the facts" in accordance with Article 11 of the DSU which provides that:[15]

> … 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 in the covered agreements.

The Appellate Body was surely correct in seeking a compromise between the extremes of *de novo* review and complete deference. *De novo* review limit the ability of domestic interest groups to exert undue influence on decision-makers in order to obtain regulations favourable to them. However, it would inappropriately constrain countries' regulatory autonomy by making panels (or their experts) the final arbiters of which regulatory response a country should have taken, or, in Trebilcock and Soloway's words, it would turn panels into a "global science court".[16] Complete deference, on the other hand, would render the panel review process meaningless and leave panels unable to control protectionist tendencies.

Yet, as Trebilcock and Soloway observe, the Appellate Body's approach of requiring a panel to make an objective assessment of the facts before it is not particularly helpful either.[17] It does not shed any light on exactly what facts are to be objectively assessed. As Prévost argues, the so-called objective standard allowed the Panel in *EC – Hormones* to essentially substitute its own judgment for that of the Member government, however incorrect its appreciation of the scientific evidence before it.[18] Further, she notes that in substituting their own opinion, panels are not obliged to give as much weight as the

---

[15] *EC/Hormones AB Report*, at para 117.
[16] Michael Trebilcock & Julie Soloway, "International Policy and Domestic Food Safety Regulation: The Case for Substantial Deference by the WTO Dispute Settlement Body Under the SPS Agreement" in Daniel Kennedy & James Southwick, eds., *The Political Economy of International Trade* (Cambridge: Cambridge University Press, 2002) at 564.
[17] *Ibid*.
[18] Denise Prévost, "What Role for the Precautionary Principle in WTO Law After *Japan-Apples*?" (2005) 2:4 Journal of Trade and Environment Studies 1 at 11.

regulating Member might desire to statements of caution by scientific experts.[19] Thus, while the objective standard is unobjectionable in itself, further clarification is required if it is actually to have any meaning and to truly prevent panels from second-guessing domestic agencies' scientific conclusions.

In *Japan – Apples*, Japan argued that in reviewing evidence under Article 2.2, the Panel was obliged to favour Japan's approach to risk and scientific evidence over the views of the experts. The Appellate Body rejected this argument, noting that it conflicted with a standard requiring objective assessment of the facts.[20] It said that the Panel enjoys a margin of discretion in assessing the value of the evidence and weight to be ascribed to it.[21] While this could be interpreted as a rejection by the Appellate Body of a deferential approach, it can also be seen as specific to the facts of the case, where Japan was in effect asking the Panel to show total deference to and effectively not to review its risk assessment at all.

Returning to hormones, the Dispute Settlement Body (DSB) adopted the Panel and Appellate Body rulings in 1998 and recommended that the EC bring its measures into compliance with WTO rules. The EC did not do so by the required date and in July 1999 the US and Canada obtained authorization from the DSB to retaliate against the EC. They subsequently introduced duties in excess of bound rates on various products from EC member States.

In the meantime, the EC commissioned 17 new scientific studies to assess the risks to human health posed by the six hormones at issue. On the basis of these studies and additional scientific information obtained, the EC's Scientific Committee on Veterinary Measures relating to Public Health (SCVPH) issued three Opinions between 1999 and 2002 concerning risks to human health posed by the hormones. Based on these Opinions, in September 2003 the EC adopted Directive 2003/74/EC which permanently banned the

---

[19] Prévost, *ibid*. at 12. See also Theofanis Christoforou, "Settlement of Science-Based Trade Disputes in the WTO: A Critical Review of the Developing Case Law in the Face of Scientific Uncertainty" (2000) 8 N.Y.U. Envtl. L.J. 622 at 645.
[20] *Japan/Apples AB Report*, *supra* note **Error! Bookmark not defined.** at para 165.
[21] *Ibid*. at para 166.

use of one of the hormones (oestradiol-17β) for animal growth promotion, and provisionally banned such use of the other five hormones while it sought more complete scientific information.

The EC stated that with introduction of the new Directive, it had brought itself into compliance with the DSB's recommendations and rulings. The US and Canada responded that the new Directive was not based on science, and they continued retaliation. In response, in November 2004 the EC requested consultations with the US claiming that the US should have removed its retaliatory measures. Canada subsequently joined those consultations. In May 2005, the EC requested composition of a panel. The case heard by this panel is known as the *US – Continued Suspension* case.

One of the key issues considered by both the Panel and the Appellate Body in *US – Continued Suspension* was whether the EC's 2003 Directive was based on a risk assessment. The EC argued that it had now concluded a comprehensive risk assessment which focused on the potential risks to human health from hormone residues in bovine meat and meat products. The Panel found, however, that the EC had not satisfied the Article 5.1 risk assessment requirement because "the scientific evidence referred to in the Opinions does not support the conclusions therein".[22] In making this determination, the Panel relied heavily on the opinions of its experts whom it had appointed to assist it in understanding and interpreting the scientific evidence.

The Panel commenced its discussion of the standard of review applicable to the determination of consistency with the SPS Agreement by recalling the Appellate Body's statement in *EC – Hormones* that Article 11 of the DSU requires panels to make an objective assessment of the matter before it, "including an objective assessment of the facts of the case".[23] It also referred to the Appellate Body's statement in the same case that a risk assessment may be based on divergent or minority views.[24] It noted that a panel is not to engage in a *de novo* review of the evidence, but also that it is "not

---

[22] *Ibid*. at para 7.578.
[23] *US – Continued Suspension Panel Report*, at para 7.412.
[24] *Ibid*. at para 7.408.

expected to refer to all statements made by the experts advising it and should be allowed a substantial margin of discretion as to which statements are useful to refer to explicitly as long as we do not deliberately disregard or distort evidence".[25] The difficulty with this statement is immediately apparent – how ought a panel decide which statements to refer to and which not to refer to? The problem inherent in SPS cases involving complex scientific evidence is that panelists are not scientific experts, and lacking specialist knowledge, it is very difficult for them to be sure that they are focusing on the most relevant statements. This is why taking an approach that essentially reconsiders the minutia of the scientific conclusions reached by domestic regulatory bodies is not an appropriate course of action for panels.

The Panel went on to explain that it had appointed six scientific experts on an individual basis so that it might "benefit from hearing the full spectrum of experts' views and thus obtain a more complete picture of the mainstream scientific opinion and of any divergent views".[26] The Panel emphasized that it was "not carrying out its own risk assessment".[27] Yet there is a problem here too. If six experts provide advice to a panel, and that advice comprises a spectrum of opinions, are we not right back to where the panel began – confronted by conflicting scientific evidence which it is incapable of fully comprehending and judging? This statement by the Panel suggests that it was doing exactly what it said it should not be doing (and what it is patently unqualified to do), namely, conducting a *de novo* review.

The Panel explained that in some cases, the experts expressed divergent opinions and that "on some occasions, we followed the majority of experts expressing concurrent views, in some others the divergence of views were such that we could not follow that approach and decided to accept the position(s) which appeared, in our view, to be the most specific in relation to the question at issue and to be best supported by arguments and evidence".[28] Again, the Panel appeared to be putting itself in the position of the respondent when it

---

[25] *Ibid*. at para 7.416.
[26] *Ibid*. at para 7.418.
[27] *Ibid*.
[28] *Ibid*. at para 7.420.

started its risk assessment process. It is listening to experts who agree with one conclusion, and experts who agree with another, and is making its own decision as to which conclusion is preferable. While the question of which position is most specific in relation to the question at issue may be an appropriate one to ask the experts, surely the question for the Panel should be whether the conclusion chosen is scientifically credible. Here, however, while the Panel emphasizes that its members are not scientists,[29] and that it is not carrying out its own risk assessment, it seems that in fact it is doing just that.

The EC appealed the Panel's decision. It claimed that the Panel had not accorded sufficient deference and argued that by following an inappropriate standard of review, the Panel had decided to "become the jury on the correct science … by picking and choosing between conflicting and contradictory opinions of the experts in an arbitrary manner".[30] In doing so, argued the EC, the Panel had impermissibly engaged in a *de novo* review of the EC's risk assessment, and failed to take into account divergent views among the experts reflecting a genuine and legitimate scientific controversy.[31] The Appellate Body agreed that a *de novo* standard of review is not appropriate. It stated that the review power of a panel is "not to determine whether the risk assessment undertaken by a WTO Member is correct, but rather to determine whether that risk assessment is supported by coherent reasoning and respectable scientific evidence and is, in this sense, objectively justifiable.[32]

The Panel had explained to the experts that their role was "*inter alia*, to present scientific issues to the Panel members in a way that could be understood by them".[33] This is unobjectionable in itself. It does not mean that either the experts or the panelists ought to go so far as to choose the conclusion they wish to believe. Rather, the role of the experts should be to focus on helping the panel determine whether or not the respondent country

---

[29] *Ibid*.
[30] European Communities - appellant's submission, at paras 239-240.
[31] European Communities - appellant's submission, at para 248.
[32] *US – Continued Suspension AB Report*, at para 590.
[33] *US – Continued Suspension Panel Report*, at para 7.420.

has complied with the SPS Agreement in terms of performing a scientifically sound risk assessment and basing their measures on the same.

Not only had the Panel erred by essentially second-guessing the EC's scientific evidence, but the Appellate Body also found that it had used the expert's opinions as the basis of a general discussion as to the relevant risks, rather than looking specifically at the evidence relied upon in the EC's risk assessment.[34] It noted, for example, that the Panel had considered the genotoxicity of oestradiol-17β in light of the EC's conclusion that a threshold could not be established for oestradiol-17β. The Appellate Body found that the Panel had focused incorrectly on examining the general acceptance of the scientific basis of the EC's risk assessment.[35] While relying on two of the scientific experts who did not find evidence of genotoxicity, the Panel did not explain how it reconciled its conclusion with the evidence of the other experts who appeared to acknowledge that the EC's assessment was based on valid scientific evidence.[36] The Appellate Body, rightly in this author's opinion, stated that the Panel was not called upon to determine whether there is general acceptance that oestradiol-17β is genotoxic *in vivo* or that it causes cancer by a genotoxic mechanism. Instead, the Appellate Body held, "the focus should have been on the evidence relied upon by the European Communities in its risk assessment".[37]

The Appellate Body concluded that "ultimately, the Panel reviewed the scientific experts' opinions and somewhat peremptorily decided what it considered to be the best science, rather than following the more limited exercise that its mandate required".[38] The Appellate Body did not specifically endorse "deference", yet its approach is undeniably more deferential than intrusive. With respect to the Article 5.1 risk assessment requirement, the Appellate Body suggested a threefold test for determining compliance. First, respectable scientific evidence, while it need not reflect the majority view within

---

[34] *US – Continued Suspension AB Report*, at para 603.
[35] *Ibid*.
[36] *US – Continued Suspension AB Report*, at para 609. For example, Dr Guttlenplan suggests that the EC's risk assessment did specifically examine the potential for adverse effects from the consumption of meat from cattle treated with oestradiol-17β, yet the Panel did not address that response or explain why it did not consider the testimony to support the EC's position. (at para 613)
[37] *US – Continued Suspension AB Report*, at para 610.
[38] *US – Continued Suspension AB Report*, at para 612.

the scientific community, must "have the necessary scientific and methodological rigour to be considered reputable science". This should be judged "according to the standards of the relevant scientific community."[39] This is a very similar approach to what this author has argued for elsewhere.[40] The Appellate Body does not detail what it means by the "relevant community", leaving some uncertainty in relation to this part of the test. The relevant community could arguably refer to the international community of scientists in a particular field of evidence as evidenced through publications in the most well-respected journals. Second, the Appellate Body said that the reasoning articulated on the basis of the scientific evidence must be objective and coherent. In other words, the conclusions drawn in the risk assessment must find sufficient support in the scientific evidence relied upon. Third, the results of the risk assessment must "sufficiently warrant" the SPS measure at issue.

The Appellate Body's requirements will clearly require panels to fall back on expert assistance. The Appellate Body clarified what panels may and may not do in terms of relying on the advice of experts. Panels *must not* consult experts in order to perform their own risk assessment and they must not test whether the experts would have done a risk assessment in the same way and would have reached the same conclusions as the risk assessor.[41] However, panels *may* seek expert assistance to: (i) identify the scientific basis of the SPS measure and to verify that this scientific basis comes from a qualified and respected source; (ii) review whether the reasoning articulated on the basis of the scientific evidence is objective and coherent, and whether the particular conclusions drawn by the Member assessing the risk find sufficient support in the evidence; (iii) clarify whether the particular conclusions drawn by the Member assessing the risk find sufficient support in the evidence; and (iv) determine whether the risk assessment "sufficiently warrants" the SPS measure.[42]

---

[39] *US – Continued Suspension AB Report*, at para 591.
[40] Tracey Epps, *International Trade and Health Protection: A Critical Assessment of the WTO's SPS Agreement* (Cheltenham: Edward Elgar, 2008).
[41] *US – Continued Suspension AB Report*, at para 592.
[42] *US – Continued Suspension AB Report*, at para 592.

The Appellate Body found that in this case, the Panel had *not* sought and used expert assistance in the appropriate manner. The Panel had summarized the experts' responses before turning to the actual risk assessment in question. The Appellate Body considered that the Panel should have looked at the EC's risk assessment first, and then asked the questions set out above.[43] Instead, the Panel had conducted a survey of the experts' advice and based its decision on whether the majority of the experts, or the opinion that was most thoroughly reasoned or specific to the question at issue, agreed with the conclusion in the EC's risk assessment. This, the Appellate Body found, was not consistent with the applicable standard of review.[44]

The Appellate Body illustrated the Panel's perverse reasoning process by referring to its analysis of the genotoxicity of oestradiol-17β. The Appellate Body suggested that the Panel's first step should have been to identify what in the EC's risk assessment was the scientific basis for the conclusions on the hormone's genotoxicity, to verify whether this scientific basis came from a respected and qualified source, and to determine whether the reasoning articulated on the basis of that scientific evidence is objective and coherent. Second, the Panel should have pursued a similar inquiry concerning the EC's conclusion that the genotoxicity of oestradiol-17β did not permit the establishment of a threshold. In this context, the Panel would have sought the experts' views as to whether the EC's conclusions could find support in the scientific evidence relied upon.[45] Rather than following this course, however, the Panel – noting that it was not in a position to evaluate the scientific data reviewed by the SCVPH – began with a survey of the experts' views on the hormone's genotoxicity. The Appellate Body noted that under the applicable standard of review, neither the Panel nor its experts were required to evaluate the correctness of the EU's risk assessment, but should have confined themselves to identifying the scientific basis for and evidence relied upon in the risk assessment, verifying that the scientific evidence comes from respected and qualified sources, and

---

[43] *US – Continued Suspension AB Report*, at para 598.
[44] *US – Continued Suspension AB Report*, at para 598.
[45] *US – Continued Suspension AB Report*, at para 601.

determining whether the reasoning articulated by the EC on the basis of the scientific evidence is objective and coherent.[46]

5. Has the Appellate Body resolved the problem?

As Trebilcock and Soloway note, if too wide a degree of deference is afforded to a Member's regulations, and any remotely plausible explanation can be offered as a rationale for a trade-restrictive health regulation, international trade rules risk being undermined.[47] On the other hand, if not enough deference is afforded, then panels would have the power, in effect, to invalidate a Member's regulations that, while trade-restricting, represent a legitimate domestic policy choice.

In determining an appropriate level of deference, two important truths must first be considered: first, the lack of objectivity of science; and second, the mixed motives of domestic decision-makers. With respect to the first point, there is a significant body of literature concerning the objectivity or otherwise of science. It is strongly arguable that pure objectivity is neither possible nor even plausible, particularly in the health and ecological sciences where uncertainties abound. The human body, animals, and plants are all complex systems, about which strong scientific assertions often break down, introducing the possibility of multiple outcomes.[48] Further, while scientific facts may exist on a neutral basis, their interpretation will often be subject to subjective values. In addition, science is vulnerable to manipulation and capture; able to be used disingenuously either to support regulations with no rational underlying scientific merit; or to question the validity of (meritorious) evidence on which regulations are based.[49] Manipulation and capture may result from various sources, including scientists

---

[46] *US – Continued Suspension AB Report*, at para 602.
[47] Michael Trebilcock & Julie Soloway, "International Policy and Domestic Food Safety Regulation: The Case for Substantial Deference by the WTO Dispute Settlement Body under the SPS Agreement" in Daniel Kennedy & James Southwick, eds., *The Political Economy of International Trade* (Cambridge: Cambridge University Press, 2002) at 541.
[48] Jeffery Atik, "Science and International Regulatory Convergence" (1996-97) 17 Nw. J. Int'l L. & Bus. 736 at 747.
[49] See William A. Kerr, "Science-based Rules of Trade - A Mantra for Some, An Anathema for Others" (2003) 4:2 The Estey Centre Journal of International Law and Trade Policy 86.

themselves, governments, employers, and industry sponsors. As Atik argues, science is vulnerable to becoming nothing more than "another political ideology explicitly directed by money and power".[50]

Yet despite these weaknesses, scientific disciplines have much to offer in providing a baseline of legitimacy for health regulations. It is arguably the best tool we have to sift out unwanted protectionism from genuine health protection measures. Without science, panels are at the mercy of competing unproven claims and perspectives. Science is capable at some level of being rationalized and assessed, even if that process is not perfect. However, assessment by panels of a country's scientific evidence must be cognizant of the possibility of its misuse by domestic regulatory authorities. Thus a rigorous (although not *de novo*) procedurally-oriented standard of review is called for under Article 5.1 in order to provide confidence that spurious scientific claims will be discovered.

With respect to the second point, numerous theorists have sought to explain what influences regulatory decision-making. This is important because if dispute panels could be certain that decision-makers always acted in the public interest, they may be justified in taking a more deferential approach than if regulators are likely to be influenced by private interests seeking regulatory rents. There is a broad distinction between 'public' interest and 'private' interest theory. Public interest theory assumes that regulation is developed in order to further the 'public interest' (such as health protection).[51] Private interest theory (such as public choice[52]), on the other hand, sees regulatory decisions as a competition for influence by various private interests.[53] The latter may result in

---

[50] Jeffery Atik, "Science and International Regulatory Convergence" (1996-97) 17 Nw. J. Int'l L. & Bus. 736 at 758.
[51] Robert Baldwin & Martin Cave, *Understanding Regulation: Theory, Strategy, and Practice* (New York: Oxford University Press, 1999) at 19. 440
[52] Public choice theory finds that legislators are motivated primarily by maximizing their chances of election or re-election. To this end, they support programs and vote for regulations that are most responsive to groups who provide financial support or politically meaningful endorsements. Nicholas Mercuro & Steven G. Medema, *Economics and the Law: From Posner to Post-Modernism* (Princeton, New Jersey: Princeton University Press, 1997) at 92.
[53] Robert Baldwin & Martin Cave, *Understanding Regulation: Theory, Strategy, and Practice* (New York: Oxford University Press, 1999) at 21.

protectionist measures that benefit local industry but violate WTO rules. Other theories of regulatory decision-making emphasize different factors. Trebilcock, for example, has written about notions of distributive justice, corrective justice, due process, and communitarianism.[54]

Another possible influencing factor is the SPS Agreement itself. Domestic regulatory agencies are constrained in their decision-making by the Agreement's rules. This discipline, (in addition to other domestic rules governing the administrative process), arguably insulates governments and administrative regulators from interest group pressure.[55] The question arises as to whether or not regulators are also constrained by the prospect of measures being the subject of complaint by another country and further, the standard of review that will be applied to that complaint. Research indicates that in the domestic context, agencies may be quite strongly impacted by the prospect of judicial review and the approach taken by courts to that review. Tiller finds for example that in the US, judicial intervention leads to agency decisions engaging in less ambitious policy change or perhaps no change from the status quo.[56] It is unlikely that domestic agencies will be as strongly affected by the approach taken by WTO panels. Extremely few domestic decisions are ever reviewed by a WTO disputes panel and it is therefore unlikely that they play any significant role in agency behaviour. It is likely that the rules themselves (the fact that SPS measures have to be justified by science and notified to other countries) have more impact than the way the rules are actually applied or compliance therewith actually reviewed.

Ultimately, no one theory can, by itself, provide a conclusive explanation for regulatory behaviour in every case. Yet, while no theory is exhaustive, each has some value. The possibility that public choice theory has some explanatory power speaks to the

---

[54] Michael J. Trebilcock, "The Choice of Governing Instrument: A Retrospective" in Pearl Eliadis, Margaret M. Hill & Michael Howlett, eds., *Designing Government* (Montreal & Kingston: McGill-Queen's University Press, 2005) 51 at 54.
[55] Stephen P. Croley, *Regulation and Public Interest: The Possibility of Good Regulatory Government* (Princeton and Oxford: Princeton University Press, 2008) at 26.
[56] Emerson H. Tiller, "Controlling Policy by Controlling Process: Judicial Influence on Regulatory Decision Making" (1998) 14:1 The Journal of Law, Economics and Organization 114.

desirability (from a trade liberalization perspective) of a more intrusive standard of review, while acknowledgement of public interest theory and the possibility that domestic agencies seek to act in compliance with administrative procedures and WTO rules augurs for a more deferential approach. How then – considering these conflicting stories – should one address the question of identifying an appropriate standard of review?

I have argued elsewhere for a relatively deferential approach to the standard of review in SPS cases, suggesting that panels should focus on procedural propriety when evaluating Members' scientific evidence.[57] To be clear, by "relatively deferential", I do not mean that panels should simply defer to the manner in which the regulating member chooses to evaluate the scientific evidence. Rather, I mean that panels should focus on reviewing the procedural propriety of the risk assessment by asking whether it has been conducted in a rigorous manner, and whether risk estimates meet a minimum threshold of scientific rationality. In other words, the review should focus on the underlying scientific merits of the manner in which the risk assessment was conducted, rather than on the actual decision reached.[58] The purpose of the examination should not be to second-guess the scientific judgments made, but to filter out sham judgments that seek to advance protectionist interests or that have no basis at all in science ("junk science").[59]

In *US – Continued Suspension*, the Appellate Body has outlined a standard of review that in essence does this; it emphasizes scientific and methodological rigour, and the objectiveness and coherency of the scientific evidence presented. This is a significant step in the right direction. It allows a certain degree of deference to domestic decisions

---

[57] Tracey Epps, *International Trade and Health Protection: A Critical Assessment of the WTO's SPS Agreement* (Cheltenham: Edward Elgar, 2008).
[58] See for a discussion of this type of review, United States Environmental Protection Agency (EPA), *Science Policy Council Handbook: Peer Review* (Washington DC: EPA, 2000) at para 3.2.1.
[59] See also Michael Trebilcock & Julie Soloway, "International Policy and Domestic Food Safety Regulation: The Case for Substantial Deference by the WTO Dispute Settlement Body under the SPS Agreement" in Daniel Kennedy & James Southwick, eds., *The Political Economy of International Trade* (Cambridge: Cambridge University Press, 2002). They argue that the purpose of requiring at least a credible minority of scientific opinion is to screen out "junk science", on the one hand, while avoiding attempts to resolve genuine scientific uncertainty or controversy on the other. Guzman argues that panels should leave the evaluation of science, decisions about risk and the relationship between science and SPS measures to domestic governments. Andrew T. Guzman, "Food Fears: Health and Safety at the WTO" (2004) 45 Va. J. Int'l L. 1 at 37.

because countries are free to reach their own scientific decisions – subject to procedural constraints around their methodological rigour and coherency. This is not total deference, but neither is it *de novo* review. Countries will be scrutinized on their agencies' procedures, but retain discretion as to the final outcome reached.

The steps outlined by the Appellate Body are thus reasonable. However, it is not clear that the next panel to be confronted with complex scientific facts will be able readily to achieve the stated standard of review. This is because the Appellate Body's steps are broad enough that panels may feel the only way to answer each question is to undertake a substantive review of the facts as they have done in the cases to date. For example, how is a panel to determine whether the reasoning articulated on the basis of the scientific evidence is "objective and coherent"? What exactly does it mean to review a country's scientific methodology? Based on the approach taken to date, it is probable that panels will address these types of questions by continuing to seek the views of different experts as to what they consider the best evidence and then choosing between them. I have argued elsewhere that the concept of peer review as utilized by scientific journals may be able to provide some guidance for panels in their task.[60] By this I mean that panels should engage experts to help them critically review the regulating country's risk assessment to see that it is technically accurate, competently performed, and properly documented. They should ask experts to review the risk assessment, including the assumptions, calculations, extrapolations, alternative interpretations, methodology, acceptance criteria, and conclusions reached. The review would thus be – as suggested above – procedurally based, focusing on the underlying scientific merits of the risk assessment, rather than on the actual decision reached.

I have also argued that panels should appoint experts as a group, rather than as individuals. The use of an expert group would enable the experts to engage in a collective, consensus-based decision-making process.[61] This would mean that panelists

---

[60] Tracey Epps, *International Trade and Health Protection: A Critical Assessment of the WTO's SPS Agreement* (Cheltenham: Edward Elgar, 2008).
[61] Theofanis Christoforou, "Settlement of Science-Based Trade Disputes in the WTO: A Critical Review of the Developing Case Law in the Face of Scientific Uncertainty" (2000) 8 N.Y.U. Envtl. L.J. 622 at 639.

will not have to evaluate widely divergent scientific views and would hopefully be less inclined to take on the role of scientific reviewers themselves.[62]

Requiring that countries have conducted a rigorous risk assessment consistent with protocols agreed by the international scientific community, and that the scientific evidence relied upon by an importing party passes a minimum threshold of rationality is deferential to a point. It is deferential in the sense that – recognizing the uncertain nature of science – it allows countries regulatory space to reach conclusions that may differ from what another country's scientists would have reached based on the same evidence. This is important because the US National Research Council has estimated that a typical risk assessment consists of about fifty separate assumptions and extrapolations meaning that many risk assessments can be neither proven nor disproved scientifically.[63] Assumptions may be risk-averse, risk-tolerant, or risk-neutral and different assumptions can alter the outcomes significantly, with different outcomes or conclusions legitimately justified based on the science.[64]

The term "deferential" in the sense that it is used here means that different outcomes may be tolerated. However, such outcomes must be based on sound scientific methods and cannot be based simply on unjustifiable preferences or methods of the regulating country. The focus on sound methodology is particularly critical given the vulnerability of science to capture by vested interests.

Regulating countries must be able to show that they have taken the appropriate steps in terms of accepted scientific methodologies. Requiring countries to document their internal processes should, in itself, enable panels to recognize protectionist measures (or at least to ensure that such measures also have a genuine basis in health protection). It

---

[62] *Ibid*.
[63] National Research Council, *Risk Assessment in the Federal Government: Managing the Process* (Washington DC: National Research Council, 1983) at 11.
[64] Mark R. Powell, *Science at EPA: Information in the Regulatory Process* (Washington DC: Resources for the Future, 1999) at 127.

should support the SPS Agreement's trade liberalization goal by requiring countries to engage in a more comprehensively rational, deliberative decision-making process.[65]

In the domestic context, Stephenson has argued that taking a "hard look" at an agency's record of decision-making provides a reviewing court with valuable information as to the likely quality of the agency's scientific, economic, social, and political analyses.[66] If an agency is able to supply the court with an explanation that appears reasonable to generalist judges, along with sufficient references to record evidence, it is presumed to be both necessary and sufficient for a reviewing court to conclude that the regulation is justified. Stephenson refers to this as the signaling theory. It is based on the notion that providing a high-quality explanation is costly to an agency – consuming time, money, and staff that could have been devoted to other things. Therefore, the quality of the agency's explanation provides valuable information to the court even if the court cannot verify the agency's substantive analysis.[67]

The theory may be relevant in the SPS context where panelists are not experts in risk assessment. A panel may justifiably uphold a domestic decision if it is accompanied by a sufficiently detailed and high-quality record. The idea is that because the agency has gone to such an effort to defend its measure, it must be of such value to the agency that it can be considered genuine. The fact that panels may not fully understand the scientific intricacies involved is not important – the record stands for itself. Of course, the question of protectionism is less the domestic context. (It is not absent altogether, however, as various interest groups may capture regulations to their benefit.) Stephenson's theory therefore requires further consideration in the international trade context. Certainly, it is only likely to hold to the extent that panels carefully review domestic procedures as argued for here. A complete record of risk assessment which shows uncertainties in the data and assumptions made may help panels by acting as a signal of the country's genuine intent to protect health.

---

[65] [cite Howse]
[66] Matthew Stephenson, *A Costly Signaling Theory of Hard Look Review* (John M. Olin Center for Law, Economics, and Business, Discussion Paper No. 539, Harvard Law School, Cambridge, MA, 2006) at 7.
[67] *Ibid*. at 11-12.

A final point should be made. Trading partners do not simply need to accept another's measure that is based on a risk that is of either extremely low probability or which has minor consequences. This will be a case where the measure may well be more trade restrictive than necessary under Article 5.6 and where there may be other reasonably available SPS measures that achieve the member's appropriate level of protection while being less restrictive to trade than the SPS measure contested.

6. Conclusion

Panels presiding over SPS-related disputes in the WTO's dispute settlement system face a difficult task. Not only do they have to contend with reviewing complex scientific evidence, but they carry the burden of ensuring the DSU contributes to an appropriate balance between the WTO's trade liberalization goals and countries' domestic regulatory autonomy. The standard of review adopted in reviewing a country's scientific evidence under the SPS Agreement is critical in delineating this balance and forms an important part of the legal context within which domestic regulatory decision-makers must operate.

The Appellate Body's decision in *US – Continued Suspension* sets a welcome tone for future panels faced with disputes over scientific evidence under the SPS Agreement. In suggesting that panels ought to require countries to demonstrate the scientific and methodological rigour of their risk assessments, the Appellate Body has moved towards an approach that requires panels to scrutinize the method by which the member reaches the decision in question, rather than focusing on the outcome of the decision. Panels should focus on reviewing the underlying scientific merits of a risk assessment's methodology, rather than attempting to second-guess its outcome. Such merits are discernable from the methods followed. Careful review of those methods is essential if panels are to fulfill the task of preventing SPS measures being used as a disguised form of protectionism. However, it should not intrude unnecessarily on countries' ability to take measures to protect human, animal or plant life or health.