The Standard of Review in International SPS Trade Disputes.  
Some New Developments

Dr Łukasz Gruszczynski

1. Introduction

The problem of the applicable standard of review in sanitary and phytosanitary (SPS) disputes has recently become an important issue in the legal discussion. The earlier research tended to concentrate on appropriateness of specific disciplines of the WTO Agreement on the Application of Sanitary and Phytosanitary Measures (SPS Agreement),\(^1\) such as the obligation of scientific justification or requirement of consistency. The problem of the standard of review was somehow unappreciated and only rarely addressed.\(^2\) The recent decision of the Appellate Body in *US/Canada - Continued Suspension of Obligations in the EC – Hormones Dispute (US/Canada – Continued Suspension)*\(^3\) and the current dispute *Australia – Measures Affecting the Importation of Apples from New Zealand (Australia – Apples)*\(^4\) will probably turn the attention of scholars to this particular issue as the standard of review is a central problem in both cases.

The paper intends to contribution to this discussion. The initial idea was to describe of what I believe is a new type of the standard of review for SPS cases that was proposed by the WTO Appellate Body in *US/Canada – Continued Suspension* and then to compare it with the standard applied by the panel in *Australia – Apples*. The aim was to check how this new standard is applied in dispute settlement practice. In other words, I wanted to analyze how the *Apples* panel has translated the general guidelines set out by the Appellate Body into specific analytical tools when examining scientific justification for Australian measures. The panel expected at first to issue its final report in July 2009. However, in June 2009, the Chairman of the panel informed the WTO Dispute Settlement Body that due to complexity of the dispute, which required further consultation with scientific experts, this deadline was not possible to be met anymore. According to the new schedule, the report was expected in January

---

\(^{1}\) Agreement on the Application of Sanitary and Phytosanitary Measures, 1867 UNTS 493, signed on 15 April 1994.

\(^{2}\) There are important exceptions to this rule, compare for example, Catherine Button, *Power to Protect. Trade, Health and World Trade Organization*, Hart Publishing, Oxford and Portland: 2004.


\(^{4}\) *Australia – Measures Affecting the Importation of Apples from New Zealand*, WT/DS367, a panel report has not yet been adopted.
2010. Another change of the date came at the beginning of 2010. This time the panel informed that it will issue its final report in May 2010. Again the deadline was not met and it seems that the report will not be publicly available before August 2010. At the same time, the number of newspapers around the world announced that the panel in its draft report (which is still confidential and only disclosed to the parties of the dispute) found in favour of New Zealand.\(^5\) Unfortunately, this knowledge does not help a lot in the analysis of the report. This is the reason why I decided to concentrate in this paper on the evolution of the standard of review under the SPS Agreement. This will be supplemented with a brief discussion as to the relevance of recent changes for the *Australia – Apples* dispute. However, I will not attempt to predict how they would be actually applied by the panel in examination of specific scientific issues.\(^6\)

This paper proceeds as following. The first part provides a brief description of the SPS Agreement and its role in the governance of international sanitary (including food) and phytosanitary safety. Part two turns to the issue of the standard of review that is applied by the WTO panels in SPS disputes. It also analyzes in more details the general guidelines that were identified by the Appellate Body in the last SPS case. Part three discusses the problem of applicable standard of review in the context of *Australia – Apples* and attempts to show those issues are related to the applicable standard of review. In particular, I will attempt to identify how different types of standard of review may affect the outcome of the dispute. The last part draws some overall conclusion as to the direction of the SPS case law with regard to the standard of review.

### 2. The SPS Agreement and its disciplines

The SPS Agreement is one of the international treaties that was adopted as a part of the World Trade Organization system. The aim of agreement is to limit the negative impact of national SPS regulations on international trade,\(^7\) while guaranteeing WTO Members with a wide margin of regulatory discretion in SPS area. Consequently, while the WTO Members are expected to observe certain requirements when introducing and maintaining their SPS measures (those which may affect international trade), they remain in principle free to establish whatever level of protection they deem appropriate. In the WTO nomenclature that right is conventionally referred to as a right to establish appropriate level of protection (ALOP). It indicates the maximum extent of SPS risk that a particular Member is ready to tolerate.


\(^{6}\) That in fact seems to be impossible. Although all the parties’ submissions are available on the internet, minutes from the meeting with scientific experts that advised the panel have not yet been published.

As it was mentioned above, the SPS Agreement is concerned with national SPS measures that may have an impact on the flow of international trade. It applies to those regulations which are directed against certain specific risks. The catalogue is closed and include risks arising from the entry, establishment or spread of pests, diseases, disease-carrying organisms or disease-causing organisms; risks for human or animal life or health arising from additives, contaminants, toxins or disease-causing organisms in foods, beverages or feedstuffs as well as human health risks from diseases carried by animals, plants or products, or from the entry, establishment or spread of pests.\(^8\)

One of the central provisions of the SPS Agreement is its requirement to base measures on international standards, guidelines and recommendations (together referred to as standards). The agreement also identifies those standard-setting bodies that are relevant in this context. This includes the Codex Alimentarius for food safety issues, the World Organization for Animal Health for animal health and the International Plant Protection Convention for phytosanitary safety. A Member that decides to fully conform to such standards benefits from the legal presumption of compatibility with the relevant requirements of the agreement.

The obligation to base measures on international standards is, however, not absolute. A WTO Members may still deviate from such standards but in such case a scientific justification is required (i.e. a measure needs to be based on scientific principles and cannot be maintained without sufficient scientific evidence).\(^9\) Thus, science is used as a kind of quasi-objective criterion for distinguishing permissible from prohibited measures. In other words, it is a proxy that helps to determine whether a measure is really necessary.\(^10\) This general obligation is translated under the SPS Agreement into specific requirement of risk assessment. In particular, Article 5.1 stipulates that Members have to ensure that 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 risk assessment techniques developed by the relevant international organizations. The agreement also enumerates elements that need to be included in such risk assessment (Articles 5.2 – 5.3). This includes, among the other things, available scientific evidence, relevant processes and production methods, relevant ecological and environmental conditions but also, in case of quarantine risks, relevant economic factors such as the potential damage in terms of loss of production or sales in the event of the entry, establishment or spread of a pest or disease.

The SPS Agreement also makes clear that a WTO Member may act even if scientific evidence is insufficient to perform adequate risk assessment (i.e. assessment that meets the requirements of the agreement). In such case, a WTO Member can adopt a provisional measure that has to be based on available pertinent information. Simultaneously, a Member is obliged to seek to obtain

\(^{8}\) Annex A(1) of the SPS Agreement.

\(^{9}\) Article 2.2 of the SPS Agreement.

\(^{10}\) As to the potential of science to perform such function see e.g., Tracey Epps, *International Trade and Health Protection: A Critical Assessment of the WTO’s SPS Agreement* (Edward Elgar, Cheltenham: 2008).
additional information necessary for more objective assessment of risk and review its SPS measure within reasonable period of time (Article 5.7).

Different set of obligations is imposed on risk management phase of national regulatory process. As it was mentioned above, WTO Members remain in principle free to adopt any level of protection they deem appropriate. In theory, this may include a zero level risk, even if costs resulting from limitation of international trade exceed expected health benefits. At the same time, the Agreement introduces the idea of consistency with respect to risk measures that adopted in different but comparable situations (e.g. the same pathogen or disease). This obligation is again not absolute as a Member may differentiate its risk responses if there is justification of such differentiation or it is not arbitrary. In addition, Members need to ensure that their measures are not more trade-restrictive than required to achieve their ALOP, however, taking into account technical and economic feasibility. In this context, the SPS Agreement recognizes two concepts that may be helpful in ensuring the least-trade restrictiveness: regionalization and equivalence. In addition measures have to be applied only to the extent necessary to protect human, animal or plant life or health (Article 2.2) and cannot arbitrarily or unjustifiably discriminate between Members where identical or similar conditions prevail, including between its own territory and that of other Members (Article 2.3).

The SPS Agreement contains number of procedural obligations relating to transparency (i.e. requirements with respect to publication of national SPS measures, notification procedure), control, inspection and approval procedures (e.g. such procedures need to be undertaken and completed without undue delay and in no less favourable manner for imported products than for like domestic products). The Agreement also establishes the Committee on Sanitary and Phytosanitary Measures, which is responsible for administration of the agreement. In particular, the Committee carries out the functions necessary to implement the provisions of the agreement and the furtherance of its objectives, in particular with respect to harmonization (Article 12).

The above general rules have been clarified over last 15 years by WTO panels and the Appellate Body in a number of SPS disputes. As it will be


12 The Agreement impose in this regard a soft obligations that only requires Members, when determining the appropriate level of SPS protection, to take into account the objective of minimizing negative trade effects (Article 5.3).

13 Regionalization under the SPS Agreement is understood as an adapted measures to the SPS characteristics of the area - whether all of a country, part of a country, or all or parts of several countries - from which the product originated and to which the product is destined, while equivalence requires to accept the SPS measure of other Members as equivalent, even if these measures differ from their own or from those used by other Members trading in the same product, if the exporting Member objectively demonstrates to the importing Member that its measures achieve the importing Member’s ALOP.

14 Annex C of the SPS Agreement.

15 Until now there were six SPS disputes relating to growth promotion hormones, salmon, apples, fruit varietaes, biotech products and again hormones.
discussed in the subsequent section, the applicable standard of review was one of the issues that proved to be particularly problematic.

3. The concept of the standard of review in the SPS case law

The standard of review is conventionally understood as a level of scrutiny that is applied by a superior body (a court or a higher administrative authority) over a decision taken by other body that is subject to review. To put it in different words, the standard of review can be defined as “the degree of deference or discretion that the court accords to legislator or regulator” or “degree of intrusiveness or invasiveness into the legislator’s or regulator’s decision-making process”.16 It determines the extent of discretionary powers of a lower body in determination of facts and meaning of law applicable in a specific case. The standard of review may vary from de novo review to full deference. Under de novo review, a superior body is able to review all kinds of determinations made by an inferior body and substitute them with its own. A fully deferential standard restricts reviewing powers of a superior body, which may only check the procedural compliance without interfering with the substantive determinations of a lower body. Of course between these two extremes there are number of less or more deferential/de novo types of review. They may appear under different names such as arbitrary and capricious standard, clearly erroneous standard, reasonable deference or de novo review.

A concept of the standard of review is common to many national jurisdictions including all major continental and Anglo-Saxon systems. In national legal context, it mainly serves as a mechanism for allocating the power between different branches of government (i.e. executive and judicial). A deferential standard favour a body that takes initial decision (e.g. executive) while de novo review introduce additional checks of the other body (e.g. judicial) However, since there is no optimal and one-fits-for-all model of distribution of powers within a state, there is also no ideal and universal standard of review. The level of intrusiveness will therefore vary from jurisdiction to jurisdiction reflecting local particularities or current preferences of society (e.g. a need to guarantee a greater oversight from courts over the activity of administrative agencies).17

On the international level, the standard of review plays a similar function. It determines, alongside the substantive obligations, a distribution of powers between national governments and international bodies. De novo standard of review transfers a power to international level at the expense of prerogatives of domestic governments. In consequence, it has an obvious impact on the allocation of power between international and national, favouring former over latter. Deferential standard of review has an opposite effect, empowering national bodies and limiting competences of international authorities. As noted by one scholar “granting greater deference to the decisions of the state is


equivalent to increasing the substantive power of the state to impact trade.”

For example, if a panel has only limited competences over re-evaluation of scientific evidence (e.g. whether growth promotion hormones in cattle increase a risk of cancer for humans consuming meat), which constitutes a basis for a national SPS measure, local authorities gain a wider regulatory freedom. They may evaluate and assess scientific data in a rather unconstrained way, reflecting local preferences and particularities. On the other hand, if a scientific backup for a measure is reviewed afresh, a measure can be unjustified if a panel comes to different conclusions than national government. This aspect (i.e. distribution of powers) was clearly recognized by the Appellate Body in *EC - Hormones* case, where it stated: “the standard of review … must reflect the balance established [in WTO law] between the jurisdictional competences conceded by the Members to the WTO and the jurisdictional competences retained by the Members for themselves.”

From the methodological point of view, one also has to distinguish between the standard of review that is applied to factual and legal determinations. The first case is concerned with review of assessment of facts, while the second one relates to legal interpretations advanced by inferior body. In addition, one may speak about the standard of review that is applied by different reviewing bodies with respect to their own determinations (e.g. if a review system is two-instances). This includes, for example, the extent of review which is applied by the Appellate Body to panel’s findings. Although, this paper is mainly interested in the first type of standard of review (i.e. factual determinations), two other categories are sometimes also mentioned.

In principle, WTO does not explicitly provide any standard of review. The only exception is the Anti-Dumping Agreement, which stipulates in its Article 17.6 (i) that a panel in its assessment of the facts of the matter, is obliged to determine whether the authorities’ establishment of the facts was proper and whether their evaluation of those facts was unbiased and objective. If the establishment of the facts was proper and the evaluation was unbiased and objective, even though the panel might have reached a different conclusion, the evaluation shall not be overturned. As far as legal determinations are concerned, Article 17.6 (ii) provides that a panel is expected to interpret the relevant provisions of the Anti-Dumping Agreement in accordance with customary rules of interpretation of public international law. In addition, where a panel finds that a relevant provision of the agreement admits of more than one permissible interpretation, it needs to find the authorities’ measure to be in conformity with the agreement if it rests upon one of those permissible interpretations. The above standards are clearly deferential ones as they (in particular the standard that is applied to factual determinations) concentrate on procedural rather than

---


substantive compliance. Consequently, they give a considerable margin of discretion to national authorities.

None of the other WTO agreement, including the SPS Agreement, contains any comparable provision. In consequence, it was for the WTO dispute settlement bodies to identify applicable standard of review (either as a general standard to be applied across different agreements or a special form that is applicable in the context of a specific agreement). At least on its face, the Appellate Body opted for the first alternative and identified Article 11 of the Understanding on Rules and Procedures Governing the Settlement of Disputes (DSU)\(^21\) as a rule that determines applicable standard of review. According to the Appellate Body, Article 11 “articulates with great succinctness but with sufficient clarity the appropriate standard of review for panels in respect of both the ascertainment of facts and the legal characterization of such facts under the relevant agreement” [emphasis added].\(^22\) On that basis, the Appellate Body characterized the applicable standard as “neither de novo review as such, nor ‘total deference’, but rather the ‘objective assessment of the facts’.\(^23\) Although this observation was made in the context of the SPS Agreement subsequent case law has confirmed its validity under other WTO agreements.\(^24\)

A closer look at the finding of the Appellate Body reveals, however, its deficiencies. The Appellate Body by describing an applicable standard of review as objective missed a possibility to provide future panels with clear guidelines. In particular, it was observed in the literature that “this broad formulation does not assist in defining an operable standard of review because any assessment of the facts, whether highly deferential, marginally deferential, or not deferential at all, can be ‘objective’.”\(^25\) The expression ‘objective’ that is used in Article 11 seems to be more concerned with guaranteeing in the context of WTO of due process rights (i.e. fairness and impartiality of a panel) rather than with determination of the applicable standard of review.\(^26\) Nevertheless, the above passage has become a point of reference for WTO dispute settlement bodies and since then it has been frequently cited as a rule that elucidates the required level of scrutiny. Not surprisingly a substantive analysis of the panel reports shows that in practice different types of review are actually used, all of them being characterized as objective assessment of facts.

---


23 Ibidem, para. 117. The Appellate Body also added that “many panels have in the past refused to undertake de novo review, wisely, since under current practice and systems, they are in any case poorly suited to engage in such a review. On the other hand, ‘total deference to the findings of the national authorities’, it has been well said, could not ensure an ‘objective assessment’ as foreseen by Article 11 of the DSU” (ibidem).

24 In practice, different agreements seem to attach different types of the standard of review.


26 Cf. discussion in Gruszczynski, supra note 11, p. 14; see also Guzman, supra note 18, p. 4.
That is also a case for the SPS Agreement. The first SPS report (EC – Hormones)\(^27\) suggested that a considerable degree of deference should have been granted to WTO Members. In particular, the Appellate Body recognized that Members were entitled to base their SPS measure not only on the mainstream (i.e. the best available) science but could also rely on minority scientific opinions.\(^28\) This obviously limits the discretion of a panel as to re-assessment of scientific evidence. The panel is not allowed to condemn a national measure only because it finds that majority of scientists holds different position from the one that was accepted on the national level. A Member is also free as to choice of ALOP and this decision is not subject of any review. Moreover, the Appellate Body also signalled that the connection, which is required between risk assessment (scientific evidence) and SPS measure, is merely a reasonable relationship that may be influenced by character of risk. Although the Appellate Body did not explain that finding in details, one may rationally assume that it gives additional margin of discretion to national authorities (e.g. risk for human life and health are related to less intrusive scrutiny).\(^29\)

At the same time, the Appellate Body disagreed with the EC to apply a deferential reasonableness standard or the standard of review that is provided by Article 17.6(i) of the Anti-Dumping Agreement.\(^30\) According to the EC, both standards would require a panel to concentrate on procedural issues (i.e. whether a procedure prescribed by particular WTO agreement was followed, without going into substantive content of a national measure).\(^31\) In this context, the Appellate Body held that the obligations provided by the SPS Agreement were substantive in the nature and entailed inquiry into the substance of a decision.

The Hormones ruling left future panels only with a rough and general guidance as to the applicable standard of review in SPS cases. What it confirmed was that the panel, although entitled to examine the merits of a national decision, had to grant a WTO Member with a degree of deference.

The subsequent case law, however, has gradually engaged in more and more intrusive assessment of scientific data that were provided as a justification for national measures. Such standard of review allowed WTO panels to assess quality, persuasive force and correctness of the national scientific determinations and substitute them with their own and in practice came really close to de novo review. In Japan – Apples case,\(^32\) the Appellate Body made clear that the panel

---

\(^{27}\) The case was concerned with the European ban on importation and marketing of beef from hormone-treated cattle. The parties disputed whether the measure was based on risk assessment and supported by sufficient scientific evidence as required by Articles 2.2 and 5.1 of the SPS Agreement.

\(^{28}\) Appellate Body report, EC – Hormones, para. 194.

\(^{29}\) Ibidem.

\(^{30}\) It is not clear from the EC appellant submission whether deferential reasonableness and Article 17.6(i) standards are the same.

\(^{31}\) EC appellant submission, para. 126.

\(^{32}\) The dispute concerned the series of Japanese measures imposed against the risk of fireblight that could be imported with apples. The US questioned the scientific rational for such a strict regime.
was not obliged to favour Japan’s approach to risk and scientific evidence over the views of its experts. According to the Appellate Body, the opposite conclusion would conflict with the standard of objective assessment of the facts.\(^\text{33}\) Consequently, the panel enjoys a margin of discretion in assessing the value of the evidence, and the weight to be ascribed to that evidence. As explained by the Appellate Body “requiring panels (...) to give precedence to the importing Member’s evaluation of scientific evidence and risk is not compatible with this well-established principle.”\(^\text{34}\) Indeed, the analysis of the panel report shows that it did substituted factual determinations of national authorities (i.e. Japanese) with its own.\(^\text{35}\)

This approach was confirmed and developed further by the panel in \textit{EC – Biotech Products} case.\(^\text{36}\) The panel conducted detailed analysis of scientific evidence which was put forward as a justification of the EC measures. It chose between competing scientific claims articulated by its experts, preferring some of them over the others.\(^\text{37}\) This is also true for the next panel report in \textit{US/Canada – Continued Suspension} dispute.\(^\text{38}\) The panel confirmed that it enjoyed a wide discretion as to the choice of evidence when making factual findings. It also added that it was not “expected to refer to all statements made by [its] experts 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.”\(^\text{39}\) Although the panel admitted that under the SPS Agreement, it was not asked to carry out its own risk assessment, it also explained that in fact its position was similar to national body that produce such assessment. This meant that the panel wanted to receive a full spectrum of expert views to form an opinion on correctness of risk assessment.\(^\text{40}\) Despite the fact that the \textit{Suspension} panel subscribed to the view that a measure may be based on scientific minority opinion, in practice it preferred a role of a final arbiter as what is considered scientific and what is not. It is worth to cite here one the passages from the panel report, which perfectly summarized its approach:

while, 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


\(^{34}\) \textit{Ibidem}, para. 167.


\(^{39}\) \textit{Ibidem}, para. 7.416.

\(^{40}\) \textit{Ibidem}, para. 7.418.
most specific in relation to the question at issue and to be best supported by arguments and evidence.\textsuperscript{41}

Overall, this approach may be described as \textit{de novo} (or \textit{quasi de novo}) standard of review.

The above direction of the SPS case law is problematic in many respects. Although adopting a fully deferential standard of review in the context of the SPS Agreement is hardly an advisable option,\textsuperscript{42} a considerable degree of deference seems to be necessary. On the one hand, giving to much discretion to national governments (\textit{e.g.} by checking only a procedural compliance) appears to go against the text of the agreement, which requires some form of substantive examination with respect to risk assessment and scientific evidence. Arguably, deferential standard would also result in marginalization of the disciplines of the agreement. Procedural requirements of the SPS Agreement remain general and underdeveloped. Consequently, they can hardly be regarded as benchmarks against which national measure can be tested and assessed. Moreover, without denying advantages of procedural checks (\textit{i.e.} increased transparency of decision-making process that may help to detect instances of interferences from rent-seeking groups), such a model will be probably leaky with many measures simply escaping the scrutiny of WTO dispute settlement bodies.\textsuperscript{43}

On the other hand, \textit{de novo} review (including the review that was effectively adopted by 	extit{Biotech} and 	extit{Suspension} panels) is also undesirable. From the pragmatic point of view, panels seem to lack sufficient scientific competence to make a judgement over complex technical and scientific issues. Although they are assisted by the experts, it is still “very difficult for them to be sure that they are focusing on the most relevant statements”\textsuperscript{44} advanced by such experts. Panellists need to decide how to interpret opinions of experts, how to assess and chose between conflicting views or what kind of interferences to make on the basis of such opinions.\textsuperscript{45} This obviously increases a risk of producing “bad” decisions that do more harm than benefit.\textsuperscript{46}

Lack of expertise is, however, not the only problem that is faced by the dispute settlement bodies when evaluating scientific evidence. As it is broadly admitted in the literature, assessment of risk is not purely scientific task and

\textsuperscript{41} \textit{Ibidem}, para. 7.420.

\textsuperscript{42} \textit{But cf.} Adrew Guzman, \textit{Dispute Resolution in SPS Cases}, in Dan Horovitz, Daniel Moulis & Debora Steger (eds.), \textit{Ten Years of WTO Dispute Settlement}, International Bar Association (2007).

\textsuperscript{43} Trebilcock & Solowy, \textit{supra} note 7, p. 541 (noting that “if too wide a degree of deference is afforded to the Member’s regulation, and any remotely plausible explanation can be offered as a rationale for trade-restrictive health and safety standards, the world trading system risks being seriously undermined with attendant global and domestic welfare losses in gains from trade”).

\textsuperscript{44} Tracey Epps, \textit{To defer or not to defer? Has the Appellate Body resolved the issue of an appropriate standard of review in SPS cases}, p. 12 (on file with the author).

\textsuperscript{45} On the problems related to limited competence of a panel in science-based disputes, \textit{see e.g.} Gruszczynski, \textit{supra} note 11, pp. 140-2 and the literature cited there.

\textsuperscript{46} \textit{Cf.} Trebilcock & Solowy who make their argument with reference to idea of consumer welfare which increase is ultimate goal of the WTO system (Trebilcock & Solowy, \textit{supra} note 7, p. 542).
depends on the specific socio-cultural conditions of a particular country.\textsuperscript{47} Such assessment inevitably involves subjective judgments of the assessors that reflect their attitudes toward particular risks, values of the relevant community in which the experts are acting or other normative elements such as required level of protection or treatment of uncertainties. Consequently, the same set of scientific data may produce entirely different risk estimates in various jurisdictions. Again the panel seem to be worse placed to make such judgements as compared to a WTO Member maintaining the measure.\textsuperscript{48} If a panel ends up imposing on the WTO Members its own vision of science, these normative and context-dependent elements will be lost, while the ultimate determination will most probably fail to produce a correct result. Although, \textit{de novo} standard of review allows avoiding Type 2 errors (allowing protectionist and “unnecessary” measure to escape a scrutiny from WTO obligations), it may also produce Type 1 errors (condemning a measure that actually protects health and safety). Is it worth to lost one statistical life over some additional trade liberalization? How to compare damage to environment caused by the invasion of foreign species with potential gains generated by trade increase? These are difficult decisions that should be address rather on national than international level. Moreover, potential costs of a mistake on the side of the WTO dispute settlement bodies seems to considerably higher in SPS disputes than a case of other trade controversies. A loss of statistical life as consequence of removing trade barrier appears to be more costly (as a matter of principle) than any damage to international trade. Another cost is generated by the problem of compliance. In case health and safety measures (which are, by their nature, very sensitive cases) the risk of non-compliance of a defendant again seems to be higher than in other more traditional trade disputes. This obviously affects the credibility of WTO dispute settlement system as well as costs of suspension of concession (which will be more frequent than in other cases).

Another factor, which speaks against \textit{de novo} review, is the impact that it may have on the length of the whole dispute settlement process. Note that one of the functions of the WTO is efficient settlement of disputes. In particular, Article 3.3. of the DSU provides that “the prompt settlement of situations in which a Member considers that any benefits accruing to it … under the covered agreements are being impaired by measures taken by another Member is essential to the effective functioning of the WTO and the maintenance of a proper balance between the rights and obligations of Members.” Elaborating on this rule Article 12.9 stipulates that as a general rule a panel is required to issue its report within six months. An extension is possible but it cannot exceed nine months.\textsuperscript{49} \textit{De novo} review, which requires detailed examination of complex scientific evidence, prolongs that process considerably as compared to disputes decided under the other agreements. The \textit{Biotech} and \textit{Suspension} cases are good examples here. Both of them lasted for years (3 and almost 4 years respectively).


\textsuperscript{48} Button, supra note 2, p. 181.

\textsuperscript{49} In practice a panel proceeding takes an average of 12 months.
The same is true for the *Australia – Apples* where the publication of the panel report has been already postponed two times while the composition of the panel dates back to 2007. One may legitimately ask whether such delay still meets the requirement of prompt settlement of a dispute as provided by the DSU. A more deferential standard under which a panel would concentrate on methodological rather than substantive issues have a potential to effectively shorten that process.

The recent decision of the Appellate Body in the *Suspension* case may be seen as a response to the above concerns. From the legal point of view, it constitutes an important shift in the WTO jurisprudence as the Appellate Body opted for more deferential standard of review, which gives an additional margin of direction to national governments in the SPS area.

The Appellate Body overturned the *Suspension* panel on many important points. One of them was the applicable standard of review for assessment of factual determinations, including scientific findings. At first, the Appellate Body reminded that it is task of a particular Member to perform risk assessment while a panel is only entitled to review it. As a consequence, a panel is not expected to determine whether a risk assessment is correct but only whether such assessment is supported by coherent reasoning and respectable scientific evidence.\(^{50}\) The Appellate Body went on to explain what steps should be taken by the panel when performing its limited task. Thus, a panel must:

a) identify the scientific basis underlying SPS measure,

b) verify that the scientific basis comes from a respected and qualified source,

c) assess whether the reasoning articulated on the basis of the scientific evidence is objective and coherent, and

d) determine whether the results of the risk assessment sufficiently warrant the SPS measure at issue.\(^{51}\)

Both mainstream science as well as minority scientific opinions may provide the basis for a measure. What is important is not a status of particular scientific claim but rather its epistemic value. This should be assessed using the methodological parameters of relevant scientific community (*i.e.* does a specific information is defendable as scientific evidence and not whether in the opinion of particular scientist it is correct). Once a scientific basis is established and its epistemic status determined, a panel needs to examine the coherence and objectivity in the interpretation of those raw data. Although the Appellate Body did not explain this particular point, one may assume, in line with the previous point, that the task of a panel is also limited here. Consequently, conclusions drawn by a WTO Member on the basis of scientific evidence has to be only checked against scientific logic (*i.e.* whether a claim is justified in the light of particular methodology that may be described as scientific, does a claim is objective or rather biased?).\(^{52}\) Finally, a panel needs to inquiry into a relationship

---

\(^{50}\) Appellate Body Report, *US – Continued Suspension*, para. 590.

\(^{51}\) *Ibidem*, para. 591.

\(^{52}\) See also, Gruszczynski, *supra* note 11, p. 144.
between those conclusions and an SPS measure in order to assess whether they support a measure. This is probably the second point where the Appellate Body was too enigmatic. It remains unclear what level of compatibility between those two elements is required here. One may only assume, in line with the existing case law, that scientific findings may support a whole range of different SPS measures. As noted by one of the SPS panels, “there may conceivably be cases where a Member … would be justified in applying (i) an SPS measure even though another Member might not decide to apply any SPS measure on the basis of the same risk assessment, or (ii) an SPS measure which is stricter than the SPS measure applied by another Member to address the same risk.”53 Nevertheless, some additional guidance as to methods, which may be used for assessing existence of sufficient relationship between a conclusion of risk assessment and an SPS measure, seems to be desirable. Last but not least, the Appellate Body warned that the consultations with the expert advisors could not seek to test whether they would have done a risk assessment in the same way as a WTO Member but only to review such assessment as to its scientific value.54 This obviously affects the consultation process that takes place between experts and a panel. In particular, the limited role of a panel determines what questions are appropriate to put to experts and what kinds of answers are relevant to the issues before a panel.55

The above approach differs considerably from the approach of the previous jurisprudence. It explicitly prohibits a panel from inquiring into a correctness of evidence and instructs to concentrate on methodological issues in order to assess epistemic value and coherence of scientific findings. A panel may equally check (in order to verify the objectivity of a process) whether authorities have collected and considered all relevant evidence and whether the evaluation of this evidence was unbiased. This standard leaves a WTO Members with a great degree of discretion as how to assess scientific data and what kind of interferences to make on their basis. It also greatly reduces a need of the panel to engage into detailed examination of scientific evidence and deciding which scientific view is better. In theory once it is established that that a particular claim is scientifically defensible, even if improbable, the task of the panel ends.

The initial reactions from academic circles to the report of the Appellate Body report were very positive. Some even claimed that the Appellate Body decided to move to fully deferential approach under which the main focus would on the process of risk assessment (probably combined with the assessment of epistemic value of evidence) rather than its outcome. In this context, a parallel was draw between the approach of the Appellate Body and that of the US Supreme Court in the Chevron case.56 I believe that such claims seem to be somehow overstated. The test provided by the Appellate Body appears to go beyond a merely procedural approach (understood broadly as also encompassing examination of evidence against certain objective epistemic criteria). Arguably

55 Trebilcock & Soloway, supra note 7, p. 57
(c) and (d) (whether the reasoning articulated on the basis of the scientific evidence is objective and coherent and whether the results of the risk assessment sufficiently warrant the SPS measure at issue) require to perform a substantive analysis (albeit deferential one). Moreover, an inquiry into underlying methodology can be also connected with different levels of scrutiny. As noted by Button, detailed methodological assessment may in practice come close to de novo review that concentrates on the substance of the evidence.\(^{57}\)

The recent dispute between Australia and New Zealand over apples will shed more light on what is actually required from a panel under the SPS Agreement. Although the report is not yet available there are some indications that the panel used rather intrusive standard of review (either it understood the instruction of the Appellate Body in a different way or decided to deeply scrutinize the methodology used for collecting and evaluation scientific evidence that were put forward by Australia). First, the length of the whole procedure suggests that the panel engaged in detailed examination of underlying science and risk assessment. As it was mentioned at the beginning of the article, the whole proceeding lasted for more almost three years while the panel justified a delay by referring to complexity of scientific issues involved in a case. Second, according to information published in press, Australia lost a case. Again this may indicate a rather intrusive review of Australian measure as from the purely procedural point of view Australia appears to follow the instructions of the SPS Agreement.

4. Australia – Apples dispute

The dispute between Australia and New Zealand over importation of apples has a long history. New Zealand is an important producer and exporter of apples. The Australian market, however, has remained closed to New Zealand apples since 1921 when Australia introduced a ban on phytosanitary grounds (following introduction and establishment of fire blight in New Zealand in 1919). The aim of Australian measures was to prevent entry of fire blight bacteria (i.e. *Erwinia amylovora*) on its territory.

In 1986 New Zealand officially requested access to the Australian market and subsequently renewed its request several times. However, a number of risk assessments conducted by Australian agencies, including the latest one from 2006 (Import Risk Analysis or the IRA 2006) concluded that apple fruits from New Zealand could provide a transmission pathway for fire blight. The IRA 2006 also identified two other pests as potentially dangerous - European canker (fungus) and apple leafcurling midge (fly). The IRA 2006 suggested that import should be allowed but prescribed a special standard operating procedure (SOP) to be agreed between Australia and New Zealand. The SOP was expected to define phytosanitary procedures for each of the pests of quarantine and the responsibilities of both parties. New Zealand decided to engage in the

\(^{57}\) Button, *supra* note 2, p. 186 (if a reviewing body goes into details of underlying methodology its task is not so different from the body that reviews substance of evidence, the only difference is that a body would concentrate on methodological issues rather than substantive).
consultation process with Australia but the SOP had not been eventually agreed upon (the parties could not reach a consensus on number of important issues). According to New Zealand, the IRA 2006 was a document of limited scientific value while the whole process of its adoption was strongly influenced by the domestic industry fearing competition from New Zealand apples producers.\(^58\) In addition, New Zealand believed that the conclusion of the SOP was unrealistic due to inflexible position of Australia.

Consequently, New Zealand decided to initiate WTO dispute settlement procedure. It claimed that Australian SPS measures (among the other things):

a) were not supported by sufficient scientific evidence as required by Article 2.2 of the SPS Agreement (i.e. measures were based on false contention that mature, symptomless apples provide a pathway for transmitting fire blight);\(^59\)

b) did not comply with the requirements of Article 5.1 of the SPS Agreement (i.e. the IRA 2006 failed to evaluate “likelihood” of entry, establishment or spread of the pests, as well as the associated potential biological and economic consequences; Australia did not evaluate the effectiveness of other measures that could be potentially applied);\(^60\)

c) did not comply with the requirements of Article 5.2 of the SPS Agreement (i.e. Australia failed to take into account available scientific evidence and the prevalence of the relevant diseases or pests).\(^61\)

In response to New Zealand allegations, Australia highlighted its special phytosanitary status (geographically isolated island with relatively short history of agricultural production) that required taking strict SPS measures.\(^62\) As to its risk assessment, Australia argued that it fully complied with the requirements of Article 5.1-5.3 of the SPS Agreement. According to Australia, New Zealand wanted to impose a uniform vision of science that was represented by mainstream research. The Australian risk assessment, on the other hand, was based on minority scientific opinions that were recognized in the WTO jurisprudence as possible justification for SPS measures. Last but not least, Australia emphasised that its risk assessment was done in an open and transparent process in which all relevant evidence was considered and assessed.

As far as the standard of review is concerned, Australia argued that the panel was obliged to grant a Member with considerable degree of discretion (specifically with respect to national assessment of scientific evidence that constituted a basis for SPS measures). In the opinion of Australia, the SPS Agreement reflected a balance of jurisdictional competences between international and national levels. This implied that certain limitations were imposed on the nature of scrutiny that panels could apply to assessment of

---

\(^{58}\) First written submission of New Zealand, WT/DS367, para. 3.16.

\(^{59}\) Ibidem, paras. 4.7 and subsequent.

\(^{60}\) Ibidem, paras. 4.151 and subsequent.

\(^{61}\) Ibidem, para. 4.412 and subsequent.

\(^{62}\) First written submission of Australia, WT/DS367, para. 38.
factual matters such as risk assessments. On the other hand, New Zealand contested that the SPS Agreement required any special deference for national assessment of risk and appreciation of scientific evidence. According to New Zealand, neither the text of the agreement nor the existing jurisprudence did support such contention.

In practice the applicable standard of review appears to be relevant with regard to the following issues:

a) **Role of the panel** – there is no agreement between the parties as to the role of the panel in the dispute. According to Australia, the task of the panel is to review the Australian risk assessment and not to evaluate scientific evidence that is behind such assessment. Consequently, the panel should not ask: “is there a risk?” but rather: “was the risk assessor’s decision objective and credible?” New Zealand submits that such limitation would not be compatible with the panel’s margin of discretion in assessing the significance of evidence.

b) **Role of the experts** - another point of disagreement is a proper role for experts. According to Australia, panel’s role cannot be alerted by virtue of its right to seek advice from experts on scientific issues (experts are not entitled to conduct de novo review of scientific evidence and they may only assist a panel in understanding the technicalities of data). In addition, a panel needs to ensure that its evaluation of evidence does not give disproportionate weight to majority scientific views that may be represented by its experts. New Zealand believes that the responses provided by the experts assisting the panel do not go beyond their mandate.

c) **Meaning of “based on”** – Australia believes that evaluation of rational relationship between risk assessment and SPS measure (“based on” as provided by Article 5.1) requires a degree of deference. In particular Australia argues that particular evaluation of risk may support a range of possible measures, and that it is up to a Member concerned to select the most appropriate measure(s) to address the risk. New Zealand does not contest that right but rather argues that Australian risk assessment does not meet requirements of the SPS Agreement in the first place.

d) **Burden of proof** – the parties disagree as what needs to be shown by New Zealand in order to establish violation of Article 5.1 and 2.2. Australia argues that applicable standard of review (i.e. deferential) requires New Zealand, WT/DS367, para. 185.

64 *E.g.* Second written submission of New Zealand, WT/DS367, para. 2.64.

65 First written submission of Australia, WT/DS367, para 199.

66 Second written submission of New Zealand, para. 2.51


68 Second written submission of Australia, WT/DS367, para. 60.

69 Second written submission of New Zealand, WT/DS367, paras. 2.674, 2.740 and 2.315.

70 First written submission of Australia, WT/DS367, para. 211.
Zealand to establish that risk assessment is so seriously flawed that the panel cannot have reasonable confidence in it.\textsuperscript{71} According to New Zealand, such a threshold is too high and finds no support in the text of the SPS Agreement or the WTO jurisprudence.\textsuperscript{72} In any case, New Zealand claims that flaws in the IRA 2006 are such that it is impossible to have any degree of confidence in the identification of risks.\textsuperscript{73}

5. Conclusions

The analysis of the WTO jurisprudence shows interesting evolution of the applicable standard of review under the SPS Agreement with respect to factual determinations. The initial rather deferential standard (e.g. \textit{EC – Hormones}) turned out to develop into intrusive examination of scientific evidence underlying national SPS measures. The panels in \textit{EC – Biotech Products} and \textit{US/Canada – Continued Suspension} engaged in the assessment of quality, persuasive force and correctness of the scientific findings advanced by the parties (in particular by a defendant). This meant that panels could substitute factual determinations (technical and scientific) with their own.\textsuperscript{74}

The above approach had an obvious impact on the allocation of power between the WTO and its Member States, favouring international level of governance over national one. \textit{De novo} standard of review empowers WTO dispute settlement bodies as it allows them to re-examine complex factual data and come to different conclusion than local authorities. However, as discussed in this paper, such an approach raises number of difficult questions as to the legitimacy of panel’s decisions (e.g. competence of a panel as a generalist to decide on complicated scientific issues, distribution of costs related to potential erroneous decisions).

The recent decision of the Appellate Body in \textit{US/Canada – Continued Suspension} arguably constitutes an important shift in the WTO jurisprudence. The Appellate Body decided to apply more deferential standard of review, which gives an additional margin of direction to national governments in the risk regulatory process. Although, it would be too far to label it as procedural one (or purely deferential), this approach definitively constitutes a departure from the previous case law. Under this new standard, the role of the panel is limited. It may only check the epistemic value of the scientific evidence and objectivity of risk assessment process (including the objectivity of a relationship between the outcome of risk assessment and an SPS measure).

\textsuperscript{71} Second written submission of Australia, WT/DS367, para. 24 (particularly noting: “in Australia’s view, the Panel’s standard of review conditions the burden of proof borne by New Zealand as the complainant, because New Zealand cannot ask the Panel to consider arguments or evidence which would require the Panel to exceed its limited mandate.”).

\textsuperscript{72} Second written submission of New Zealand, WT/DS367, para. 2.67.

\textsuperscript{73} \textit{Ibidem}, para. 2.301.

\textsuperscript{74} As to the reasons which could explain such developments of the SPS case law, see Lukasz Gruszczynski, \textit{Science and the SPS Agreement – Changing Standard of International Trade Dispute Adjudication?} Leuven Working Paper (2010, forthcoming).
The panel report in *Australia – Apples* will be the first one to elaborate on the general instruction provided by the Appellate Body and apply it in practice. Thus, the future report appears to be of particular importance as it will shed a light on the future direction of the SPS case law. This will also allow clarifying the precise division of competences between Member States and supervisory bodies at the WTO.